

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



**STIC-Biotech/ChemLib**

40482

**From:** Rao, Manjunath N.  
**Sent:** Wednesday, April 18, 2001 3:50 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** Sequence search request for 09/262,126

**From:** Manjunath N. Rao  
Art Unit 1652, Room 10D04  
Phone: 306-5681

**Date:** 4-18-01

Please search the following as soon as possible for application with serial number **09/262,126**

SEQ ID NO.: 1 against all nucleic acid databases including issued patents database and pending application database and provide a **print of all** results.

SEQ ID NO.: 2-4 against all protein databases including issued patents database and pending application database and provide a **print of all** results.

**If you have any questions please call me at the above phone number.**

Thanks

Manjunath N. Rao  
Art Unit 1652  
Room CM1, 10D04  
Phone 306-5681

*Edward Hart  
Technical Info Specialist  
STIC / Biotech  
CM1 12C14 Tel: 305-9203*

**MANJUNATH N. RAO**  
Patent Examiner  
AU 1652, CM1 10D04  
703-306-5681





## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

## STAFF USE ONLY

## Type of Search

## Vendors and cost where applicable

Searcher: _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) <u>2</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Queste/Orbit _____
Date Searcher Picked Up: <u>4/19/01</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>7/23/01</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>03</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____



This search was run on the machines we identify as Compugen1, abss03, abss04, or abss05. The databases on these machines are exactly the same as those on the other machines. However, for searches run on the listed machines, two sets of results are now generated when the Pending Nucleic Acid and/or Pending Amino Acid databases are searched. The Pending databases have been split into two parts to reduce the amount of time required for their daily update.

Searches run against the Nucleic Acid Pending database will produce two sets of results, with the extensions **.nrpm** and **.nrpn**

Searches run against the Amino Acid Pending database will produce two sets of results, with the extensions **.rapm** and **.rapn**

*Because they contain data that is confidential, the results of Pending database searches should not be left in the case.*



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2001, 06:17:56 ; Search time 3508.21 Seconds

(without alignments)  
4504.279 Million cell updates/sec

Title: us-09-262-126c-1

Perfect score: 2794

Sequence: 1 gatgggaacagacagat.....cggtaaaagtaagaaaa 2794

Scoring table: IDENTITY\_NUC

Searched: 1283235 segs, 3078962356 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_om:\*  
9: gb\_ov:\*  
10: gb\_pat1:\*  
11: gb\_pat2:\*  
12: gb\_ph:\*  
13: gb\_pl1:\*  
14: gb\_pl2:\*  
15: gb\_pl3:\*  
16: gb\_pl4:\*  
17: gb\_pl5:\*  
18: em\_ba1:\*  
19: em\_ba2:\*  
20: em\_ba3:\*  
21: em\_ba4:\*  
22: em\_ba5:\*  
23: em\_ba6:\*  
24: em\_ba7:\*  
25: em\_ba8:\*  
26: em\_ba9:\*  
27: em\_ba10:\*  
28: em\_ba11:\*  
29: em\_ba12:\*  
30: em\_ba13:\*  
31: em\_ba14:\*  
32: em\_ba15:\*  
33: em\_ba16:\*  
34: em\_ba17:\*  
35: em\_ba18:\*  
36: em\_ba19:\*  
37: em\_ba20:\*  
38: em\_ba21:\*  
39: em\_ba22:\*  
40: em\_ba23:\*  
41: em\_ba24:\*  
42: em\_ba25:\*  
43: em\_ba26:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_v13:\*  
61: gb\_v14:\*  
62: gb\_v15:\*  
63: gb\_v16:\*  
64: gb\_v17:\*  
65: gb\_v18:\*  
66: gb\_v19:\*  
67: gb\_v20:\*  
68: gb\_v21:\*  
69: gb\_v22:\*  
70: gb\_v23:\*  
71: gb\_v24:\*  
72: gb\_v25:\*  
73: gb\_v26:\*  
74: gb\_v27:\*  
75: gb\_v28:\*  
76: gb\_v29:\*  
77: gb\_v30:\*  
78: gb\_v31:\*  
79: gb\_v32:\*  
80: gb\_v33:\*  
81: gb\_v34:\*  
82: gb\_v35:\*  
83: gb\_v36:\*  
84: gb\_v37:\*  
85: gb\_v38:\*  
86: gb\_v39:\*  
87: gb\_v40:\*  
88: gb\_v41:\*  
89: gb\_v42:\*  
90: gb\_v43:\*  
91: gb\_v44:\*  
92: gb\_v45:\*  
93: gb\_v46:\*  
94: gb\_v47:\*  
95: gb\_v48:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2786.4	99.7	4464	A38128	A38128 Sequence 8
2	2786.4	99.7	4464	A38129	A38129 Sequence 9
3	2786.4	99.7	4464	AR000474	AR000474 Sequence
4	2786.4	99.7	4464	AR000475	AR000475 Sequence
5	2786.4	99.7	4464	AR044548	AR044548 Sequence
6	2786.4	99.7	4464	AR044549	AR044549 Sequence
7	2786.4	99.7	4464	AR098239	AR098239 Sequence
8	2786.4	99.7	4464	AR098240	AR098240 Sequence
9	2786.4	99.7	4464	189311	189311 Sequence 8
10	2786.4	99.7	4464	189312	189312 Sequence 8
11	2786.4	99.7	4464	189323	189323 Sequence 8

12	2786.4	99.7	4464	10	189324	189324	Sequence 9
13	2786.4	99.7	4464	10	193619	193619	Sequence 8
14	2786.4	99.7	4464	10	193620	193620	Sequence 9
15	2776.4	99.4	2784	9	A38130	A38130	Sequence 10
16	2776.4	99.4	2784	9	AR000476	AR000476	Sequence
17	2776.4	99.4	2784	9	AR044550	AR044550	Sequence
18	2776.4	99.4	2784	9	AR098241	AR098241	Sequence
19	2776.4	99.4	2784	10	189313	189313	Sequence 10
20	2776.4	99.4	2784	10	189325	189325	Sequence 10
21	2776.4	99.4	2784	10	193621	193621	Sequence 10
22	1823.6	65.3	3934	10	E14086	E14086	Sequence
23	894.8	32.0	3972	2	BANCONHA	BANCONHA	Sequence
24	321.4	11.5	2371	10	E03513	E03513	Sequence
25	301	10.8	2500	2	BTU67061	BTU67061	Sequence
26	300	10.7	303250	2	AP001518	AP001518	Sequence
27	283	10.1	2777	3	TMPLA	TMPLA	Sequence
28	282.4	10.1	11008	1	AE001821	AE001821	Sequence
29	269.6	9.6	8128	1	AF096862	AF096862	Sequence
30	267.8	9.6	215640	2	BSU80016	BSU80016	Sequence
31	267.8	9.6	220050	1	AF008220	AF008220	Sequence
32	254.6	9.1	2744	1	AF060205	AF060205	Sequence
33	221.8	7.9	4983	2	CDCPULA	CDCPULA	Sequence
34	170.6	6.1	248254	84	SPNEU1903	SPNEU1903	Sequence
35	76.2	2.7	1993	12	A93710	A93710	Sequence
36	74.6	2.7	3261	12	AF080567	AF080567	Sequence
37	73.6	2.6	784	9	A98396	A98396	Sequence
38	72.2	2.6	2646	10	I20414	I20414	Sequence
39	72.2	2.6	2982	10	I20409	I20409	Sequence
40	72.2	2.6	2988	10	I20408	I20408	Sequence
41	69	2.5	3322	14	RICFUJ1	RICFUJ1	Sequence
42	68.2	2.4	3437	9	AS2184	AS2184	Sequence
43	68.2	2.4	3437	14	SOPULSPO	SOPULSPO	Sequence
44	68.2	2.4	3439	2	KPN5DE	KPN5DE	Sequence
45	65	2.3	2718	13	AF252635	AF252635	Sequence

ALIGNMENTS

PAT 05-MAR-1997

RESULT 1  
LOCUS A38128 4464 bp DNA  
DEFINITION Sequence 8 from Patent EP0605040.  
ACCESSION A38128  
VERSION A38128.1 GI:2294739  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Deweer, F. and Amory, A.  
TITLE Pullulanase, micro-organisms producing the same, method for  
JOURNAL Preparation thereof as well as its use  
Patent: EP 0605040-A 8 06-JUL-1994;  
SOLVAY (BE)  
COMMENT Other publication JP 6217770 940809  
Other publication CN 2112028 940629  
Other publication CN 1090325 940803  
Other publication AU 5275993 940707  
Other publication FI 935900 940629  
Other publication BE 1007723 951010  
Other publication BE 1007313 950516  
Other publication BE 1006483 940913.  
Location/Qualifiers  
1. 4464  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN

Query Match 99.7%; Score 2786.4; DB 9; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	gagggagacacgacacacatctgtccacatcttctgcccctgtgtgtatcaacct	60						
DB	1250	GATGGGAACACGACACGATCTGTCCACTATTTTGGCCCTGCTGTGATATCAACT	1309						
QY	61	tggagctatgtatgtgcccacaaagcgggtgtggctgtatcaatctcaacacg	120						
DB	1310	TGGAGCTATGATGATGCGCCAAAAGACGAGGTGGGCTGATTCGATTCATCAACCG	1369						
QY	121	gtgacaccttltgagcgtgtgtcgaagtcgtatattccagaagaaacaaatcagtgga	180						
DB	1370	GCTGACCTTTTGGAGCTGTGGCAAGTCGATATTCAGCAAAACCCAACTCAGCTGCA	1429						
QY	181	atactcgtcgcacccaagaattgacacaaagatgtgacgcgtgacgcgtacatagatt	240						
DB	1430	ATTATGCTTGCACTCAAGTGTGACCAAAAGATGTGACCCCTGACCCCTCATATGATTA	1489						
QY	241	agcaaaaggaatgtgagtggtgtgtgtgaaggaagaaacgcaaatlttataatga	300						
DB	1490	AGCAAAAGGAATGTGAGTGTGTGAAAGGAAGCAAAATTTTATATGA	1579						
QY	301	gatgctgaagatgtgacgttaaacccgcgtgaagcagccttatagatgtctcaacac	360						
DB	1550	GATGCTGAGATGTGACGCTTAACCCGCTGTAGCAACCTTATATGCTTCAACCG	1609						
QY	361	gtgctgttaaaacttaagcagcgttlaactctltggggaaggnnaagcagccttacg	420						
DB	1610	GTGCTGTAACTTACCGACCGCTTAACCTTGTGGGAAGNNNAACGGCTTACGGTT	1669						
QY	421	ctggaagacacgaacaaataagatattccagtgacatctgtgaagatgtcaagctt	480						
DB	1670	CATGACACACACGCAATTAAGATATTCACATCTGTGAAGATGCAAGCTTGT	1729						
QY	481	caagatgtaacgcgtgtgtgtgacaggtaccttccaaataatlttltgaggttccg	540						
DB	1730	CAAGATGTAAACCGCTGTGTGGCAGGTACCTTCCAAATATTTTGGAGCTTCGATTGG	1789						
QY	541	gcacctgataatcacagctacttataaaaggttgacaaatctctcaatctca	600						
DB	1790	GCACCTGATATACACAGTACTTATTAATAAAGTGACTCAATCTCTATTCATTCCTCA	1849						
QY	601	ggagatcttccctgaaggaacacaaataaagtccttaaatgtagctgtgaat	660						
DB	1850	GGAGATCTTCTGAAAGAACTACCAATTAAGTGTAAATGATACCTGCAATTAAT	1909						
QY	661	ccgagttaccatcttgacacatatttaacagttccctccgaggggtgacacgctac	720						
DB	1910	CCGAGTTACCCATCTGCACATTAATTTTACATCTCCCTCCGCGGTGACACAGTCTACT	1969						
QY	721	tttcgtataatccgtccactcatgacgtctatgaacaaatlaataatctcaatg	780						
DB	1970	TTTTCGTATATTCGCTCAGTCACTGAGTGTATGACCAATTAATCTTAATGCGAT	2029						
QY	781	ttcaagtagaagcgggtgttaaaagcgtcgtgtgacggttactctcaagggaaat	840						
DB	2030	TTACAAATGAAGCGGGGTTAAACGATCTCGTGCGGTACTCTAGGGGAAGATCCA	2089						
QY	841	gagtgagcactctgtccatctcaacagatgggtcacaagcgaagcaggtgtatcc	900						
DB	2090	GATGTGACCATCTCTGTCATTCATCAACAGATGCTATGACCAAGCGATATCT	2149						
QY	901	cgtaaatgtcttaattatcacagctactatctcaggaatgactcttggaaatac	960						
DB	2150	CGTAATGTCTTATATTCATCAGTACTACTATTCAGAGATGATCTTGGGAATACCTAT	2209						
QY	961	acacagaaacaaac	1020						
DB	2210	ACACAGAAACCAACCTTTAAAGTGTGGGACCAACTTCTACTCAAGAAATGTCTT	2269						
QY	1021	cttatgacagtgcaaggggttctgtacacaaatcttactatgacagcgtcgagcat	1080						
DB	2270	CTTATGACAGTGCACAGCGTTCTGTACAAAAATCTTACTATGACGATCGGCGCAT	2329						

QY 1081 ggtgtgtggaagcaacgggttaatacaaaaccttgaaaatttgatataatgaagta 1140  
 |||||  
 Db 2330 ggtgtgtggaagcaacgggttaatacaaaaccttgaaaatttgatataatgaagta 2389  
 |||||  
 QY 1141 acaagccaagctctaccccgaaacggctgttgatctctatgcaacgcgcatgacccaat 1200  
 |||||  
 Db 2390 acaagccaagctctaccccgaaacggctgttgatctctatgcaacgcgcatgacccaat 2449  
 |||||  
 QY 1201 ggaagcaagcagatgattgtagacctggtctaaacagatccctgctgctgagacaagtat 1260  
 |||||  
 Db 2450 ggaagcaagcagatgattgtagacctggtctaaacagatccctgctgctgagacaagtat 2509  
 |||||  
 QY 1261 aaacatatacgcacaaagaataagaagaatgagatcatatgaaatgagatgctgctgac 1320  
 |||||  
 Db 2510 aaacatatacgcacaaagaataagaagaatgagatcatatgaaatgagatgctgctgac 2569  
 |||||  
 QY 1321 ttctccatgaccttaattctggtatgaaataaagaaggagatattggtctctacaga 1380  
 |||||  
 Db 2570 ttctccatgaccttaattctggtatgaaataaagaaggagatattggtctctacaga 2629  
 |||||  
 QY 1381 aaagaaacaaagggccctgacaaagtaaaagcgggagatagattccttaaaacaaacttg 1440  
 |||||  
 Db 2630 aaagaaacaaagggccctgacaaagtaaaagcgggagatagattccttaaaacaaacttg 2689  
 |||||  
 QY 1441 attactatgctcagctatagctgctgttctgcatctaacagtgctgagtaaaactgacaa 1500  
 |||||  
 Db 2690 attactatgctcagctatagctgctgttctgcatctaacagtgctgagtaaaactgacaa 2749  
 |||||  
 QY 1501 acccaagataattggtgtatagacccctcgcaactatgattgttccctgaaggagagatgct 1560  
 |||||  
 Db 2750 acccaagataattggtgtatagacccctcgcaactatgattgttccctgaaggagagatgct 2809  
 |||||  
 QY 1561 acaaatgcaatgtaatgctgctgataaaaggtttaagaaatggttcttccactcat 1620  
 |||||  
 Db 2810 acaaatgcaatgtaatgctgctgataaaaggtttaagaaatggttcttccactcat 2869  
 |||||  
 QY 1621 cgtgaacacattggtgttaacaatgagatgctgataaaacacacacacacacacacacac 1680  
 |||||  
 Db 2870 cgtgaacacattggtgttaacaatgagatgctgataaaacacacacacacacacacacac 2929  
 |||||  
 QY 1681 tctgactcgttaaatgctgctgataaaatgctgataaaatgctgataaaatgctgataaaat 1740  
 |||||  
 Db 2930 tctgactcgttaaatgctgctgataaaatgctgataaaatgctgataaaatgctgataaaat 2989  
 |||||  
 QY 1741 ccaacggaatcaggtactgaaatgaaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 1800  
 |||||  
 Db 2990 ccaacggaatcaggtactgaaatgaaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 3049  
 |||||  
 QY 1801 attatgattcccttaagatattggtgcaatgagatcatatgacggtcttccgtttgac 1860  
 |||||  
 Db 3050 attatgattcccttaagatattggtgcaatgagatcatatgacggtcttccgtttgac 3109  
 |||||  
 QY 1861 ttaatgagcgtgctgtaaaagacacagatgctcaaaagctgctgaggttcaatgctat 1920  
 |||||  
 Db 3110 ttaatgagcgtgctgtaaaagacacagatgctcaaaagctgctgaggttcaatgctat 3169  
 |||||  
 QY 1921 aatcaagaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 1980  
 |||||  
 Db 3170 aatcaagaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 3229  
 |||||  
 QY 1981 gatcagctctgacaacaaagagctcaaaagcagatgaggaatgaggtgttcaatgcaat 2040  
 |||||  
 Db 3230 gatcagctctgacaacaaagagctcaaaagcagatgaggaatgaggtgttcaatgcaat 3289  
 |||||  
 QY 2041 ttacgaagaagcgttgaggaagcaatgcttctgatttcttctgctcaaggttttcgacagat 2100  
 |||||  
 Db 3290 ttacgaagaagcgttgaggaagcaatgcttctgatttcttctgctcaaggttttcgacagat 3349  
 |||||  
 QY 2101 gcaacagcgttaacgcatgcaatgaagaatgagcgttgaggaaggaatgaatgaatgaatgaat 2160  
 |||||  
 Db 3350 gcaacagcgttaacgcatgcaatgaagaatgagcgttgaggaaggaatgaatgaatgaatgaat 3409  
 |||||

QY 2161 tcttaccaggtgagacaatgaactatgcaacagtcagatgaataactacaccccttgagac 2220  
 |||||  
 Db 3410 tcttaccaggtgagacaatgaactatgcaacagtcagatgaataactacaccccttgagac 3469  
 |||||  
 QY 2221 aaataagcccttaagcaatcttaatgattcccgaaagcgatcgatgaataatgagatgaac 2280  
 |||||  
 Db 3470 aaataagcccttaagcaatcttaatgattcccgaaagcgatcgatgaataatgagatgaac 3529  
 |||||  
 QY 2281 gcaacagcgttgattatgaccccaacagcgcttccatcatatgcaagggggaagaatg 2340  
 |||||  
 Db 3530 gcaacagcgttgattatgaccccaacagcgcttccatcatatgcaagggggaagaatg 3589  
 |||||  
 QY 2341 ctctgtanaaaagcggcaacgaatagttataatgcaagcgatgctgcaatgatt 2400  
 |||||  
 Db 3590 ctctgtanaaaagcggcaacgaatagttataatgcaagcgatgctgcaatgatt 3649  
 |||||  
 QY 2401 gattgagcaggaagaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 2460  
 |||||  
 Db 3650 gattgagcaggaagaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 3709  
 |||||  
 QY 2461 ctctgttctgatacccaacgcttccgcatgacgacagcaatgaaatgaatgaacacac 2520  
 |||||  
 Db 3710 ctctgttctgatacccaacgcttccgcatgacgacagcaatgaaatgaatgaacacac 3769  
 |||||  
 QY 2521 caatccctaataagtcacagacagacagtgctatgaaatgaatgaatgaatgaatgaatgaat 2580  
 |||||  
 Db 3770 caatccctaataagtcacagacagacagtgctatgaaatgaatgaatgaatgaatgaatgaat 3829  
 |||||  
 QY 2581 gacaaatgggggaatatactgattgttataacccaataaaacgtaagcaacatcaat 2640  
 |||||  
 Db 3830 gacaaatgggggaatatactgattgttataacccaataaaacgtaagcaacatcaat 3889  
 |||||  
 QY 2641 ttgcgaagcgggaaatgggcaatgctacgaagcggtaaggaatgaatgaatgaatgaatgaat 2700  
 |||||  
 Db 3890 ttgcgaagcgggaaatgggcaatgctacgaagcggtaaggaatgaatgaatgaatgaatgaat 3949  
 |||||  
 QY 2701 ggtcaagcagaggaagtgctcaagtaacaggtatatactatgagatgaatgaatgaatgaat 2760  
 |||||  
 Db 3950 ggtcaagcagaggaagtgctcaagtaacaggtatatactatgagatgaatgaatgaatgaat 4009  
 |||||  
 QY 2761 gtaagcccaacgagcgttaaaagtaatagaana 2794  
 |||||  
 Db 4010 gtaagcccaacgagcgttaaaagtaatagaana 4043  
 |||||  
 RESULT 2  
 A38129  
 LOCUS A38129 4464 bp DNA  
 DEFINITION Sequence 9 from Patent EP0605040.  
 ACCESSION A38129  
 VERSION A38129.1 GI:2294740  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 unidentified.  
 unclassified.  
 unclassified.  
 1 (bases 1 to 4464)  
 REFERENCE  
 Dwee, P. and Amory, A.  
 Pullulanase, micro-organisms producing the same, method for  
 preparation thereof as well as its use  
 Patent: EP 0605040-A 9 06-JUL-1994;  
 SOLVAY (BE)  
 COMMENT  
 Other publication JP 6217770 940809  
 Other publication CA 2112028 940629  
 Other publication CN 1090325 940803  
 Other publication AU 5275993 940707  
 Other publication FI 935900 940629  
 Other publication BE 1007723 951010  
 Other publication BE 1007313 950516  
 Other publication BE 1006483 940913.  
 Location/Qualifiers  
 1..4464  
 /organism="unidentified"  
 /db\_xref="taxon:32644"

BASE COUNT	1355 a	855 c	1039 g	1187 t	28 others
ORIGIN					

Query Match	99.7%	Score 2786.4;	DB 9;	Length 4464;
Best Local Similarly	100.0%	Pred. No. 0;		
Matches 2793;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	ga	gggaaacagacaaacagatcatlgtlccaatatluttlgcgcgtcgglygatatacaacct	60
Db	1250	gatgggaacacgacacacgacatcttgcacatttttggccctgcgtgatattatcaccct	1309	
QY	61	tgga	ctctatgtagctgtgcccacaaagacgaggtggtggccctgaaatcagaatctcaacaacg	120
Db	1310	tggagtctatggaatgtggcccaaaaagacgaggtggggcctgaattaccatttcattacaccg	1368	
QY	121	gclga	ctclttltagagcblgtlgaaglytclgaatctccagaaacccaagtcgaagtaaga	180
Db	1370	cctgcactcttttttgagagcgtttgcacagtcgtgatatttccagaaacccaagtcagtaga	1429	
QY	181	atltac	gtlgtgcactcaaaagatltagaccacaaagatgtgaagcctggaacgctacatagaatla	240
Db	1430	atttatcgttccactccactcaaaagttagaccacaaagatgtgaagcctggaacgctacataga	1489	
QY	241	agcaaa	ggaaagaaatgaggtgtgtgtctgtatgaagaaagacacgaatcttctataatgaanaa	300
Db	1490	agcaaa	ggaaatgaggtgtgtgtctgtatgaagaaagaaacgcaaatTTTTTATATATGAAAA	1549
QY	301	ga	tlctlgaagatlgacgtcacaacccgctlgtlaagcaacgctatltlgaatgtctcaaacag	360
Db	1550	gatcctgcagatgtgcacgcttAAACCCCGCTGAACCAACGCTTATTTAGATCTTCAAAACAG	1609	
QY	361	gtgc	tggttlaaaactltagccagccgtltaaacctctgggaagagmmmaagggcctttacggt	420
Db	1610	gttcgctgttAAACTTAAACCAACCGCTTAACTCTTGGGGAAGNNNAAGCGGCTTAAAGGTT	1668	
QY	421	catga	cgcacacagacaaatlaagatatctccagtgaacatctgtlaagaaatgtcaagctcttgt	480
Db	1670	catgacgcacacagcaaaatlaagatatctccagtgaacatctgtlaagaaatgtcaagctcttgt	1729	
QY	481	caaga	tlgtlaaacgctgtlcttggcgaagtlacctctcaacatatlttttgaagtlccgaattgg	540
Db	1730	caagatgttAAACCGCTGTTTTGGCAGGTAACTTCAACAAATATTTTGGAGGTTCCCAATGG	1789	
QY	541	gcaac	gtlaatactcaagtaactlataaanaagtgaaatcaaaactctcaatactlccca	600
Db	1790	gcacctgatatatcacaactacttTATTAAAAAGGTACTAACAACTCTCATTAATTCTCA	1849	
QY	601	ggaga	ctctccctgaagaaactcaacatatlaaagtgtccttaaatgaatgaatcgtgaataat	660
Db	1850	ggaaatcttCTCTGAAAGAAATCTCAAAATATAAAGTGGCTTTAAATGATGTGGAATAT	1909	
QY	661	ccga	gttaccactlgtgaacaactlataaagtlccctgcgcgggggtggaagaatcaact	720
Db	1910	ccgagtttAAACCATCTTGACAACATTTAAATTAACAGTCCCTCCGCCGCCGTCCACACCTCACT	1969	
QY	721	lttct	gtatatlctcgtlccactlcaatgcaatctlaagaacaaatlaataacaaatgaatgcgat	780
Db	1970	TTTTCGTATATTCGTGCTCCACTCATATGCAGTGTATAGACAAATTAATATCTTAATCGGAT	2029	
QY	781	ctaac	agtaagaagcgggtgtlaaagcggatctcgtlgaacggttactctgaaggtgaagatcca	840
Db	2030	TTTAACTAGTAAAGCGGGGTTTAAACGAGATCTGTAGCGGTACTCTAGGGGAAGATCCA	2089	
QY	841	ga	tltaggccaactctgtlccatctcaaacagatlgactlcaaggcaaaagcaagtlgaact	900
Db	2090	gatctgacccctTAATCTCTCAATTCAAAACGAGATCCCTATACGCAAAACGAGTGTATTA	2149	
QY	901	cgtaa	atgtctlaatctacacagtaactaactlctcagggagatlgacatcttgggaataacctat	960
Db	2150	CGTAAATGGCTTATTCATACAGTACTACTATTCAGAGATGATGATCTTGGGAATACCTAT	2209	
QY	961	acacaga	aaagcaaacacctttaaagtlctgggaccaactctactcaagtlaaagtlttct	1020

Db	2210	ACACGAAAGCACACACTTTAAAGTCTGGCGACCACTTCTACCTCAAGTAAATGTTCTT	2269
Qy	1021	cttaatgaagatgaaacgggtctctgtaacaaaatcgtatacctatgacgcatcggccat	1080
Db	2270	CTTTATGACACTGCACACGGGTTCTGTAACAAAATGTAACATATGACGCGATCGGGCAT	2329
Qy	1081	gctgtgtggaaagaaacggttaatacaaaccttgaaatctgtattacatgtatgaagta	1140
Db	2330	GGTGTGTGGAAAGCAAGGTTAACTCAAAACCTTGAAAATGGTATTACATGTATGAAGTA	2389
Qy	1141	acaggccaaagcctctacccgaagcgtctgtacccctatgacaactgcaatctgacccaat	1200
Db	2390	ACAGGCCAAAGCCTACCCGACGGCGTTGATCCTTATGCAACTGCGATTGCCACCAAT	2449
Qy	1201	ggaaacgaagacatgattgtgagaccgtgtcaaaacagatccctgtgtcgtgaaacagtat	1260
Db	2450	GGAAAGCAAGGCATGATTTGTGACCTGGCTAAAACAGATCCCTGCGCGAACAGTAT	2509
Qy	1261	aaacatactacgcgaagaatatatgaagatgagtgacatcatagaaatgatatccgtgac	1320
Db	2510	AAACATATTACGCCAAGGAATATAGAAAGATGAGAGTACTATGAAATGATTCGGTATC	2569
Qy	1321	cttcccatcgaccctaaatctcggatgataaaataaagggaagtaattgtgtctctacgaa	1380
Db	2570	TTTTTCCATTGGACCTTAATTCGGGATGAAAAAATAAGCGACATATTGGCTCTTACAGA	2629
Qy	1381	aaaggaaacaaaggccctgcgaacgaatgaagaacggggaatagattccctaaacaaactgg	1440
Db	2630	AAAGCAACAAAAGCCCTGACACACGTAAACCGGGATAGATTCTCTTAAACAACTTGG	2689
Qy	1441	attaactcatgtcagctataatgacctcttcgcgaatcacaacagttcgtatgaaactgaca	1500
Db	2690	ATTACTCATGTTGACGCTTATGCTGTTTGGATCTTACACAGTGTGATGAACATGATCA	2749
Qy	1501	accgaagtaattgagggttatatgaacctcgaaactatgatagtctccgaaaggacgtatgct	1560
Db	2750	ACCCAGAGTAATTTGGGGTTATGACCCCTCGCACTATGATGTTCTGAAAGCGCAGATGCT	2809
Qy	1561	acaaatcgcaatgtaatagtcgcgtataaaagatctaaagaaatggtctcttcactcat	1620
Db	2810	ACAAATGCGAATGGTAATGTCGCTATMAAAGACTTTAAGGAATGGTCTTTCACCTCAT	2869
Qy	1621	cgtagaacatctgggtttaacaatgatatgtctataatcatacctctgcgaagcaatc	1680
Db	2870	CCTGAACATTTGGGGTTAACTAGGAAGTGTCTTAAATCAATACCTTGGCCACGCAATTC	2929
Qy	1681	tcgtacctcgataaaatctgataccgaatatctatccgtacagatgatacaggtatacta	1740
Db	2930	TCTGACTTCGATTAATTTGTAACCGAATATTATTACCGTACATGATGACAGTAAATTATA	2999
Qy	1741	ccaacggaatcgaatcagaaatgaaatctgcaangcmaaaagccaaatggtctcaaaaatt	1800
Db	2990	CCAAACGATTCAGTACGTACGTGAAGAAATTTGCANGCAGAAAGCCAAATGGTCCAAAATTT	3049
Qy	1801	attatgtctcccttaagattcttgatcaatgagtatacatattgaaggtctccgtcttgac	1860
Db	3050	ATTATTGATTCCTTAAGTATTGGGTCATAGATATCATATTGACGGCTCCGTTTGAC	3109
Qy	1861	ttaatgacgtcgtcttgaaagaacagatgtccaaagcgtccgtcggaactcagatcatt	1920
Db	3110	TTTAAATGGCGCTGTTGGAAAAAGACAGATGTCCAAAGTCCCTCGAGCTTATGCTATT	3169
Qy	1921	aatcagaatattgacattcaacgltgagcaatggaacgggttgaaactctgcacatgcagat	1980
Db	3170	AATCCAGGAATTTGCATTTTACGGTGAGCCATGACCGGATGGAACCTTGACACTGCCAGAT	3229
Qy	1981	gatacagctctcgaaaaaaggagctcaaaaaagatgaggtgaggggttttaatgaacat	2040
Db	3230	GATCAGCTTCTGACAAAAAGAGCTCAAAAAAGCATGGGATGGCGGTGTTTAAATGACAT	3289
Qy	2041	ttaacgaacgcgtctggaacgcaatgtcttgaattcttcgcgcgaaggttttgcgaacgt	2100



Db 3290 TTACGAACCGCTTGACGCGCAATGCTTTGATTTCTCCGCTCAAGGTTTCCGACAGCT 3349  
Qy 2101 gcaacagccttaactgataacttaagaatggcgttgaggagagatataacttaac 2160  
Db 3350 GCAACAGGCTTAAGTGAATGAATGAGGCGTTGAGGGAGATTAATGACTTTAC 3409  
Qy 2161 tcttaacagctgagacaacttaactatgcaagatcaactaactaactccttgagac 2220  
Db 3410 TCTTCACCGAGTGAGCAATTAATGATCTCACAGTATGATTAACCTTGGGAC 3469  
Qy 2221 aaatagcccttaagaactcctaactatgatacgaagagatcgatataaactgaactc 2280  
Db 3470 AAATAGGCTTAAGCAATTCCTAATGATTCGGAAGCGGATTCGATTAATGATTAACCT 3529  
Qy 2281 gcaacagccttaactgataactcacaagcgcttccatcattgacaagcgaggagaaatg 2340  
Db 3530 GCACAGAGAGTTGTTATGACTCCACAGCGCTTCATTCATGCAAGCGGGGCAAAATG 3589  
Qy 2341 ctctgtanaaaagcgagcaacagcaactatataatgcaagcgatcggtcaatgatt 2400  
Db 3590 CTTCGTANAAAAGCGGCAACGCAATAGTTATATGACAGCGGATCGGTCMAATGATT 3649  
Qy 2401 gattgagacagaaagcctaactatcgaatgtttcaactatataagcggtlaactaac 2460  
Db 3650 GATTGAGCGCAAGCAATTCATATGATGTTTCAACTATTATAGCGGCTAATCCAC 3709  
Qy 2461 ctctgcttgaatcacacagccttccgcatgacagcctaactaactaactaactcactc 2520  
Db 3710 CTTCGTCTGTATCACCAGCCTTCGCAATGAGACAGCTAATGAATCAATGCAACCTC 3769  
Qy 2521 caattcctaactatgctcagagaaacagctgacctaactaactaactaactaactaacta 2580  
Db 3770 CAATTCCTAATATAGTCCAGAGAACACAGTGGCTATGAAATTCATGATCATGTTAATAA 3829  
Qy 2581 gacaaatggggaataatattgttgttttaatacccaataaactgagcaactaact 2640  
Db 3830 GACAATATGGGAATATATGTTGTTTATATACCAATAAATCTGTGCAACCAATCAAT 3889  
Qy 2641 ttgcccagcgggaaatgggcaactaactcactaagcggtaagtgaggaaatccaccct 2700  
Db 3890 TTGCGGACGGGAATGGGCAATCAATGCTAGCGGTAAGTGAGAAATCCACCTT 3949  
Qy 2701 ggtcaagcagagagagatgttccaaagtaaccaggtatatactatgataactcactaagag 2760  
Db 3950 GGTCAAGCAGAGGAGAGTGTCCAACTACAGGTATATCATATGATGATCTTCATCAAGAG 4009  
Qy 2761 gtaagcccaagccagcgttaaaaaagtaataagaaa 2794  
Db 4010 GTAAGCCAGACCGTAAATAAGTAATGAAAA 4043

RESULT 3  
AR000474  
LOCUS AR000474 4464 bp DNA PAT 04-DEC-1998  
DEFINITION Sequence 8 from patent US 5736375.  
ACCESSION AR000474  
VERSION AR000474.1 GI:3963005  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Dwyer, P. and Amory, A.  
TITLE Expression system for novel pullulanase  
JOURNAL Patent: US 5736375-A 8 07-Apr-1998;  
FEATURES  
source 1..4464  
/organism="unknown"

BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN

Query Match 99.7% Score 2786.4; DB 9; Length 4464;

Best Local Similarity 100.0%: Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 gatggaacacagacagatcattgccaactatttgcgcgtcgtgattatcaact 60  
Db 1250 GATGGGAACACGACACCAATCTATTCTCAGTATTTTGGCTCGTGATATATCAACT 1309  
Qy 61 tggagctatagatgtggccaaagacgaggttggtggaatagattcaatcaacg 120  
Db 1310 TGGAGCTATGATGTGGCCAAAGACGAGGTGGGGCTGAATATGATTTCAACCG 1369  
Qy 121 gctgactcttttggagcgtgtgcaagtgtctatattcagaagaccagtcagtaga 180  
Db 1370 GCTGACTCTTTTGGAGCTGTTTCCAGTGTGATTTCCAGGAACCAAGTCAGTAGGA 1429  
Qy 181 attatcgttgcactcaagatggaccaaagatgagagcgtgacgcgtacataatga 240  
Db 1430 ATTATGCTTCCGACATCAAGATTGACCAAGATGAGCGGTGACCGCTACATGATTTTA 1489  
Qy 241 agcaagagaaatgaggtgtgctgtgagaggaacagccaaattttatataatgaaaa 300  
Db 1490 AGCAAGGAATGAGGTGTGCTGTAGAGGAACAGCAATTTTATATGAAAAA 1549  
Qy 301 gatgctgagatgcagcctaaccgctgaagcaacgttatttagatgctcaaacag 360  
Db 1550 GATGCTGAGATGACAGCTAAACCCGCTGAACCAACGCTTATGATCTTCAAAACAG 1609  
Qy 361 gtgctggttaaaacttaagccagcgttactcttggggaaggnnaagcgctttagt 420  
Db 1610 GTGCTGCTTAAACTTAAGCCAGCGCTTAACCTTTGGGGAAGNNNAAGCGGCTTACGGTT 1669  
Qy 421 catgacacacagcaataaagatatccagtgacacactgttgaagatgcaactgtgt 480  
Db 1670 CATGACGACACGCAATATAGATATTCAGTACATGTAAGATGCAAGTCTTGCT 1729  
Qy 481 caagatgttaaccgctgttttggcaggtaccttccaacatatcttggaggtccgattg 540  
Db 1730 CAAGATGTAAACCGCTGTTTGGCAGGTACCTTCAACATATTTTGGAGGTTCCCATTTGG 1789  
Qy 541 gcaactgataactcaagtaacttatttaaaaaaggtgaactaactcactaacttca 600  
Db 1790 GCACCTGATTAATCAAGTACTTTTATTAATAAAGGTGACTTAACAATCTTATCAATTTCTCA 1849  
Qy 601 ggaagatcttctcgaaggaactaactaataaagtggcttaactatagatgctgaataat 660  
Db 1850 GGAGATCTTCCGTAAGGAATCTACCAATATTAAGTGGCTTAAATGATAGCTGGAAATAT 1909  
Qy 661 ccgagttaccacatcgacaacatlaactaagtcctcgtccggtgacacgtcact 720  
Db 1910 CCGAGTTACCCATCTGACAAACATTAATTAACAGTCCCTGCGGCGGTGCACACGCTACT 1969  
Qy 721 ttttggatattcgttccactatgcaactatgacacaaatlaactaactaactcgtggt 780  
Db 1970 TTTTGTATATTTCCGTCCTCATGTCAGTCTATGACAAATTAATTAATCTTATGCGGAT 2029  
Qy 781 ttcaagatgaagcgggtttaaaagatctcgtgaaggttaccttaagggaagatcca 840  
Db 2030 TTCAAGATGAAGAAGCGGGTTAAACGATCTCGTACCGGTTACTTAAGGAAGATCCA 2089  
Qy 841 gatgtgagcactactctgtcacttcaaacagatggtcatalcaagcgaaggtgatacct 900  
Db 2090 GATGTGAGCCATATCTGTGTCATTAACAGATGGCTATACAGCAAGCAGGTGATACCT 2149  
Qy 901 cgtatgttcttaattatcacaagatcactatcagaagatgagatcttggagatactat 960  
Db 2150 CGTAAATGCTTAAATTCATCAAGTACTATTCAGGAGATCAATTTGGGAATATCTTAT 2209  
Qy 961 acaacgaagcaaacacttaagctgtggcaacacactcactcaagtaaatgtctt 1020  
Db 2210 ACACAGAAAGCAACAACTTTAAAGTCTGGGACCAACTTCTTACTCAATTAATGTCTT 2269  
Qy 1021 cttatgacagtgcaacgggttctgttaacaaaatcgtactatgacggtcgtggtcact 1080

Db 2270 CTTTATACAGTGCAGGGTTCTGTACAAAAATCGTACCTATGACGCATCGGCCAT 2399  
Qy 1081 ggtgtgtggaagaaagcgttataacaaaccttgaaatctgtatctacatgtatgagta 1140  
Db 2230 GGTGTGTGGAAGAACGCTTAATCAAACTTCGAAATTTGGTATATCATGTATGAGSTA 2389  
Qy 1141 acggggccaaagctctaccggaagggcgtgtgtatcccttaacatgtcgatctgcaaat 1200  
Db 2290 ACAGGCCAACGCTCTACCCGACACCGCTGTATCCCTTATACACCTGCATTCACCAAT 2449  
Qy 1201 ggaacgagagcgtatctgtggaacctgtgcaaacagatccctgtgctggaacgtgat 1260  
Db 2450 GGAACGAGAGCATGATGTGTGACCTGGCTAATAACAGATCTGCTGGCTGGACAGATGAT 2509  
Qy 1261 aaacatactacgcaaaagaaatataagaatgtatgtatctatgaatgtatgtgtgtgtac 1320  
Db 2510 AAACATATTACGCCCAAGAAATATAGAACATGAGCTCATCTATGAATGCAATGTCCTGAC 2569  
Qy 1321 ttttccatltgaccttaattctgggtatgtataaaatataaaggagaatattgtgtctctaaagaa 1380  
Db 2570 TTTTCCATTTGACCTTAATTTGGCGTATGAAAAATAAAGGGAATTTTGGCTCTTACAGAA 2629  
Qy 1381 aaagaaacaaagggcccttgcaaacgttaagacggtggaatagatcccttaaaacacgtggg 1440  
Db 2630 AAAGGAACAAAGGCCCTTGACAAAGTAAGACGGGATACATTCTTAAACAACTTGGG 2689  
Qy 1441 actactacgttcaagcttatgtctgtcttcgcatacaacgtgtcgaatgaatctatcca 1500  
Db 2690 ATTACTATGTTCAGCTTATGCGCTGTATGCGCATCTAACAGTGCATGAAACGATGCCA 2749  
Qy 1501 acccaagaataattgggttctgagacctcgcaactatgattcttcgaaggcgtatgct 1560  
Db 2750 ACCCAAGATTAATTTGGGTTATGACCCGTGCAACATATGATCTCTGTAAGGCACTATGCT 2809  
Qy 1561 acaaatctgcaatgttatactgtctataaaagagcttaagaaatgtgtctcttccatccat 1620  
Db 2810 ACAATCTGCAATGTTATGCTCGTATAAAGAGTTAAGGAATGTTCTTTCATCTCCAT 2869  
Qy 1621 cgtgaacacatctgggtgtatacaatgtatgtctgtctataacatcccttgcacgcaatc 1680  
Db 2870 CGTCAACACATTTGGGTTTAACTGATGATGTTCTATATCATATCCCTTGCACCAATTC 2929  
Qy 1681 tctgaccttgataaaatctgtacagaaatatactacgtatccgagatccggtaatatata 1740  
Db 2930 TCTGACCTTCGATAAATTTGACAGAAATTTATTACCTGACGATGATGCAAGTAATTTATA 2989  
Qy 1741 ccaacgacatcagctacgtgaatctgaatctgcangcnaaagcccaatgtctcaaaaatc 1800  
Db 2990 CCACGACATCAGGTACGTGAAATGAAATTCGANCNCAAGGCCAATGTTCAAAAATTT 3049  
Qy 1801 actatgactcccttaagatctgggtcaatgtatgatactatgaaggctccgtttgac 1860  
Db 3050 ATTATTTATTTCCCTTAAGTATTGGGTCAATGAGTATCATATTGACCGCTTCCGTTTTCGAC 3109  
Qy 1861 ttaaattgacgtcgtcttgaaagaaagacagatgccaagctccctcggaagcttcatgcatc 1920  
Db 3110 TTAATTTGCGCTGCTTGAAGAAAGACACATGTCCAACCTGCTCGAGCTTCACTCATTT 3169  
Qy 1921 aatccagaaatctgcaactctacgtgtgacacatgtgacggtgtgaaacctctgacatgcagat 1980  
Db 3170 AATCCAGGAATTTGCACTTTACGCTGAGCCATGAGCGGGTGAACCTCTGCACTCCGAGAT 3229  
Qy 1981 gatcagctcttgacaaaagagatcnaaaagacatggagatgacggtgtcttaatgacaat 2040  
Db 3230 GATCAGCTTCTGACAAAAGACGCTCAAAAAGCCATGCGAGTACGGGCTTTTAAATGACAAT 3289  
Qy 2041 ttacgaagacgcttggagacggaatgtctgtatctcgcgtcaaggcttttggaagagc 2100  
Db 3290 TTACGAACCGCTTGGACGCGCAATGCTTTGATTTCTTCCCTCAGGCTTTTGGACAGGT 3349  
Qy 2101 gcaacaggtcttaactgtatgtcaatgaatgtggtgtgagggagatcaatgtactaac 2160  
Db 3350 GCAACAGGCTTAAGTATGATGAATTAAGATGGGCTGAGGGAGATATTAAATGACTTTACC 3409

Qy 2161 tcttccacaaagtgagacaatlaactatgtatcacaagtgatcatataactacaccccttggag 2220  
Db 3410 TCTTCACACAGGTGAGACAATTAATCATGTGTACAAAGCATCATTAATACCTTTGGAC 3469  
Qy 2221 aaaaatagcccttaagcaatccttaatgtatctcgaagcggatcggatlaaatgtatgaatc 2280  
Db 3470 AAAATAGCCCTTAAGCAATCTTAATGATTCGAAAGCCGATGATGATTAATAAGATGAATC 3529  
Qy 2281 gcaacagctgtctatgtacctccaaagcgttccatctcatgtcaaggcgggaagaaatg 2340  
Db 3530 GCACAGCAATGTTATGACTCTCAAGGCGCTTCAATTCATGCAAGCGGGGAAGAAATG 3389  
Qy 2341 ctctgtanaaaagcggcaacagcaatagttataatgtcaagcgaatgcgtcaatgtatc 2400  
Db 3590 CTTCGTANAAAAGCGCGCAACGACAATATGATTAATGACAGCGATGCGTCAATGACTTT 3649  
Qy 2401 gatttggagcaggaagctcaatctccagatgttcttaacatctatagcgggtcaatccac 2460  
Db 3650 GATTGGAGCAGGAAGCTCAATATCCAGATGTTTCAACATATATAGCGGGCTAATCCAC 3709  
Qy 2461 ctctgtctgtatccaccagccttcgcatagtacgacagctaatgaatcaatagccactc 2520  
Db 3710 CTTCGTCTTGTATCACCCAGCTTCCGCATGACACACACTATGAAATCAATAGCACCTC 3769  
Qy 2521 caatctctaaatagctcagagaacacagtggtcctatgaatlaactgtatcaatgtat 2580  
Db 3770 CAATTCCTTAATAGTCCAGAGAACACAGTGGCCATGAATTAACGTATGTTAATAAAA 3829  
Qy 2581 gacaaatggggaatataatctatgttattatatacccaataaaactgtagaacatcaat 2640  
Db 3830 GACAAATGGGGAATATATCATGTTGTTATTAACCAATTAACCTGAGCAACCATCAAT 3889  
Qy 2641 ttcgcaagcgggaatgtggaatcaatctatgtctacagcgtgaagtaagaaatccaccc 2700  
Db 3890 TTGCGCAGCGGGAATGTGGCAATCAATGCTACAGCGGTAAAGTAGAGATCCACCTT 3949  
Qy 2701 ggtcaagcagaagggaaatgttccaagtacaggtatatactatgtatgactcttcaagag 2760  
Db 3950 GGTCAACACAGAGCAAGTGTCAAGTACAGGTATATCATGATGATGATCAATCAAGAG 4009  
Qy 2761 gtaagcccaagcagcggtaaaagtaataaagaa 2794  
Db 4010 GTAAGCCAGACCGGTAAAGATTAAGAAA 4043

RESULT 4  
AR000475  
LOCUS AR000475 4464 bp DNA PAT 04-DEC-1998  
DEFINITION Sequence 9 from patent US 5736375.  
ACCESSION AR000475  
VERSION AR000475.1 GI:3963006  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Deweer,P. and Amory,A.  
TITLE Expression system for novel pullulanase  
JOURNAL Patent: US 5736375-A 9 07-Apr-1998;  
FEATURES  
source 1..4464  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN

Query Match 99.7%; Score 2786.4; DB 9; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 gtagggaaacgnaaagcatatgtccactatttgcctctgtgtgtatcaacct 60  
|||||

Dh 1250 GATGGGAACGACAGCATATTGTCCATATTTTCCCTGCTGGATTAATCAACCT 1309  
Qy 61 tgaagctatgagatgagccaaagacggaagtgaggctgataacatcaacacg 120  
Dh 1310 TCGAGTGTATGATGTGGCCAAAAGACGAGGTGGGGTGAATACGATTTCATCAACCG 1369  
Qy 121 gctgaactcttggagctgttgcaagtgataltccagaagaacccaagtcaglaaga 180  
Dh 1370 GCTACTCTTTTGGACGCTTGGCAAGTCTGATATTTCCAGCAAAACCCCAAGTCAGTAGGA 1429  
Qy 181 attatcgttcgcaactcaagatltgacaaagatgtgagcgtgacgtatccgtatagatta 240  
Dh 1430 ATATTCGTTCCACTCAAGATTGGACCAAGATGTGAGCGCTGACCGCTACATAGATTTA 1489  
Qy 241 agcaaaagaaatgagtggtggtctgtgaagaaacagccaatcttccataatgaaaaa 300  
Dh 1490 AGCAAGAAGAAATGAGGTGTGGCTTGTAGACGAAMACCCCAATTTTATATATGAAAA 1549  
Qy 301 gatgctgaagatgacgtctaaacccgctglaagcaacgcttatltagatgctcaaacag 360  
Dh 1550 GATGCTGAGATGACGCTTAAACCCGCTGTAAAGCAACGCTTATTTAGATGCTTCAAAACAG 1609  
Qy 361 gtgctgtgtaacttaagcaacgcttaactcttggggaagmmaagcgcttaacgctt 420  
Dh 1610 GTGCTGTTAACTTAGCCACCGCTTAACCTTGGGGAAGGNNNAAGCGCTTAAAGGTT 1669  
Qy 421 catgacgacacagaataaagataatccagtgacatctgtgaagatgacagtcgtgt 480  
Dh 1670 CATGACGACACAGCAAAATGAAGATATTCAGTGAACATCTGTGAAGATGCAAGTCTTGCT 1729  
Qy 481 caaagttaacccgctgttctggcaggtacccccaacatlttggaggtccagttgg 540  
Dh 1730 CAAGATGTAAACCGGTGTTTGGCAGGTACCTTCCAAACATATTTTGGAGGTCCGATTGG 1789  
Qy 541 gcaactgttaactcaagacttlttaaaaagtgactaaacatctatcaatccca 600  
Dh 1790 GCACCTGTATATCAACAGTACTTTATTAANAAGTGACTAACAACTCTATCAATTTCA 1849  
Qy 601 ggaactctccctgaagaaataccaataaagtgctttaaagatagctggaataat 660  
Dh 1850 GGAGATCTTCTGGAAGGAATACCAATATTAAGTGGCTTTAAAGATAGCTGGAAATAT 1909  
Qy 661 ccgagttacccatctgacacacatlaaaltlaacagtcctgcgcgcggtgacacgctca 720  
Dh 1910 CCGAGTTACCCATCTGACAAACATTAATTAAACAGTCCGCGCGGTCACACGTCACCT 1969  
Qy 721 ttctcgtaattcgttcacacatgacgtctatgacacaaatlaaatccctaactcgat 780  
Dh 1970 TTTTCGTATATTCGCTCCACTCATGACGTATGACACAAATTAATTAATGCGGAT 2029  
Qy 781 ttacaagtagaagcggggtlaaaaacggaatcgtgaacggttaactctlaaggggaagatca 840  
Dh 2030 TTACAAAGTAGAAGGGGGGTTAAAAACGGATCTCGTACGGTTACTCTAAGGGAAGATCCA 2089  
Qy 841 gatgtgaagcaatctctgtccatltcaaacagaatggtatcaaggcaaggtgatact 900  
Dh 2090 GATGTGACCCATCTCTCTCATTTCAACAGATGCTATCAGGCAACAGGTGATACCT 2149  
Qy 901 cgttaatggttaattatcaacagtaactatltcagagatgagatcttgggaatacctat 960  
Dh 2150 CGTAAATGCTTAATTCATCAAGTACTACTACTTTCAGAGATGATCTTGGGAATACCTAT 2209  
Qy 961 acacagaagcaacacacctttaaagctctggtgacacacactctactcaagtaatgtct 1020  
Dh 2210 ACACAGAAGCAACAACTTTAAAGTCTGGGACCAACCTTCTACTCAAGTAAATGTTCT 2269  
Qy 1021 cttatgtcaatgtcaaggggtctgttaacaaaaatcgttaactatgaaggtacgtggccat 1080  
Dh 2270 CTTATATGCACTGCAACGGGTTCTGTAAACAAAATCCTACATATCAGGCACTGGGCAT 2329  
Qy 1081 ggtctgtggaagcaaggtttaaataaaaccttgaataatgtatataatgataagta 1140  
Dh 2330 GGTGTGTGGGAAGCAACGGTTAATCAAAACCTTGAAATTTGATTAATGATGAGTA 2389

Qy 1141 aagagccaaaggtctctaccggaacggtctgtatccctlaagcaactgcatgtcaccaaat 1200  
Dh 2390 ACAGGCCAAAGGCTTACCCGAACGGCTGTGATCTTTATGCAATGCGATTCCACCAAT 2449  
Qy 1201 ggaacgaagagcaatgtgtgagccttggctlaaacagatcctgtctgtgaaagtgat 1260  
Dh 2450 GGAAAGCAAGGCGATATGTGTGACCTGGCTAAACAGATCCCTGCTGGGTCGAACAGTAT 2509  
Qy 1261 aaacatattacgcaaaaataatagaatgtgcatctatgaatgtatccgtgac 1320  
Dh 2510 AAACATATTAGCCCAAGAAATATGAAAGTGAAGTCAATCTTAAGAAATGATGTCCGTGAC 2569  
Qy 1321 ttctcatgtacccttaattccggtlaagaaataaagggaaatlttggcttlaagaa 1380  
Dh 2570 TTTTCATTTGACCCCTAATTTCCGGTATGAAAAATTAAGGAATATTTGGCTCTTACAGAA 2629  
Qy 1381 aaaggaacaaagggccctgacaaaglaaagacggggaatagatltcctlaaacaacttgg 1440  
Dh 2630 AAAGGAACAAAGGGCCCTGACAAAGTAAAGAGGGGATGATTCCTTAAACAACTTGAG 2689  
Qy 1441 attactatgttcagcttatgctgttcttgatataacagtgatgataaactatcca 1500  
Dh 2690 ATTACTCATGTACCTTATGCTGTGTTTCGATCTAACAGTGTGATGAACCTAATCCA 2749  
Qy 1501 acccaagataattgggtlatalgaccctcgcaactatgatlcttcgaaaggaagatgct 1560  
Dh 2750 ACCCAAGATTAATTTGGGTTATGACCCCTGCCAATGATGCTTCGAAAGCGATATGCT 2809  
Qy 1561 acaaatgcaatlgtaatgctgtgtaaaaaggttlaaagaaatggttcttcaactcaat 1620  
Dh 2810 ACAAAATGCAATGTAATGCTGTATTAAGAGTTAAAGGAATGTTCTTCACTCCAT 2869  
Qy 1621 cgtgaacacatgggttlaagatgagtgtgctataacataactcttggcagcaaac 1680  
Dh 2870 GGTGAACACATTTGGGTTAACATGATGTGTCTTAATCAATCTTGGCCACGCAAAAC 2929  
Qy 1681 tctgaactcgaataaattgtacagaataatltacccgaagatgacgaataatla 1740  
Dh 2930 TCTGACTTGCAATAAATGTATACAGAAATTTATACCGTACGATGACGATTAATTA 2989  
Qy 1741 ccaacggaatcaggtctggaataatgaaatgcaagcngaaaggccaaatgttcaaaaaat 1800  
Dh 2990 CCMAAGGATCAAGTACTGGAATGAAATTTGCAAGCNGAAAGCCAAATGTTCAAAAATTT 3049  
Qy 1801 attatgtatccttaagatgtgggtcaatgagatcgaatgaacggtcttcggttggc 1860  
Dh 3050 ATTATTTGATTCCTTAAATGATTTGGGTCAATGATATCATATTGACGGCTTCGTTTGAC 3109  
Qy 1861 ttaatggtcgtgcttgaagaaagacagatgtlccaaagctgcgtcggagcttcatgtat 1920  
Dh 3110 TTAATGGCGCTCTTGGAAAGACAGATGTCCAAAGCTGCTCGAGCTTCATGCTAT 3169  
Qy 1921 aatccaagaatltgacatltgaagtgagcaatggaaggtggaacotctgcactgccagat 1980  
Dh 3170 AATCCAGGAATTTGCACTTTACCGTGAAGCATGAGCGGGGGAACCTCTGACATGCCAGAT 3229  
Qy 1981 gatcagctctgacaaaagagctcaaaaagacatgggggtagcggtgttlaaagaat 2040  
Dh 3230 GATCAGCTTCTACAAAGGAAGACTCAAAAAGGCAATGGAGTACGGGTGTTAATGCAAT 3289  
Qy 2041 ttacgaacagcgtltgagcggcaatgtlcttgaatcttccgcctcaaggtlttgcgaaggt 2100  
Dh 3290 TTACGAAGAGCGCTTGAGACGCAATGTCTTGAATTCCTCCGTCGAAGGTTTTCGACAGAT 3349  
Qy 2101 gcaacaggtctaacgtgaatgaagaaatggtcgttgaagggagatlaaataacttacc 2160  
Dh 3350 GCMAAGGCTTAATGATGCAATTAAGATGCGTTGAAGGGAATTAATTAATGACTTAC 3409  
Qy 2161 tcttcaacaggtgagaaataatcaatgtcacaaatcaatgaataactaacaccccttgggac 2220  
Dh 3410 TCTTCAACCAAGGTGAGACATTAATTAATGTCAACAGTCAATGATTAACCTTTGGGAC 3469



QY 1201 ggaacgagagcatgtatgttgagccttgcttaaacagatcctgtgcttgagcaagtgat 1260  
DB 2450 ggaacgagagcatgtatgttgagccttgcttaaacagatcctgtgcttgagcaagtgat 2509  
QY 1261 aaacatatacgcgaagaataatagaaatgagtcacatcctatgaatgatgtccgtgac 1320  
DB 2510 aaacatatacgcgaagaataatagaaatgagtcacatcctatgaatgatgtccgtgac 2569  
QY 1321 ttcttcattgaccttaactcctggtatgaaataaagaaggaatcttggccttaacaga 1380  
DB 2570 ttcttcattgaccttaactcctggtatgaaataaagaaggaatcttggccttaacaga 2629  
QY 1381 aaagaacaaaggccctgacaaaglaaagcgggagatagatccttaaacacttgag 1440  
DB 2630 aaagaacaaaggccctgacaaaglaaagcgggagatagatccttaaacacttgag 2689  
QY 1441 attactatgttcaagcttatgtctgttttcgcatcctaaagagtgatgaaactgatcca 1500  
DB 2690 attactatgttcaagcttatgtctgttttcgcatcctaaagagtgatgaaactgatcca 2749  
QY 1501 acccaagaataatgggtatgatacctcgcacataatgattcctcgaaggagcatgct 1560  
DB 2750 acccaagaataatgggtatgatacctcgcacataatgattcctcgaaggagcatgct 2809  
QY 1561 acaaatgcgaatgtaatgtctgataaaagaagttlaaggaatggcttcttactccat 1620  
DB 2810 acaaatgcgaatgtaatgtctgataaaagaagttlaaggaatggcttcttactccat 2869  
QY 1621 cgtgaacacattgggtatgaacgtatgttctataccttaacacttggcagcaatc 1680  
DB 2870 cgtgaacacattgggtatgaacgtatgttctataccttaacacttggcagcaatc 2929  
QY 1681 tctgactcgtataaataatgtacacagaatacttaacccgtacatgacaggtatata 1740  
DB 2930 tctgactcgtataaataatgtacacagaatacttaacccgtacatgacaggtatata 2989  
QY 1741 ccaacgagatcaggtacgtgaaatgaatgacngcngaaaggccaatggttcaaaaatt 1800  
DB 2990 ccaacgagatcaggtacgtgaaatgaatgacngcngaaaggccaatggttcaaaaatt 3049  
QY 1801 attatgattcccttaaatatgagtgatgaatgataatgaagggcttccgtttgac 1860  
DB 3050 attatgattcccttaaatatgagtgatgaatgataatgaagggcttccgtttgac 3109  
QY 1861 ttaatgagcgtgcttgagaaagacagatgctcaaaagctgcctcggagcttactatc 1920  
DB 3110 ttaatgagcgtgcttgagaaagacagatgctcaaaagctgcctcggagcttactatc 3169  
QY 1921 aatcagaagaatgacattcaggtgagccaagagcgggtggaactcgtacatgccaag 1980  
DB 3170 aatcagaagaatgacattcaggtgagccaagagcgggtggaactcgtacatgccaag 3229  
QY 1981 gatcagctctgagcaaaaggaactaaaggaatgagtgagtggtgttaataacat 2040  
DB 3230 gatcagctctgagcaaaaggaactaaaggaatgagtgagtggtgttaataacat 3289  
QY 2041 ttaacaaacgcttgagcagcaatgcttctgattcctccgctcaaggttttgcaaggt 2100  
DB 3280 ttaacaaacgcttgagcagcaatgcttctgattcctccgctcaaggttttgcaaggt 3349  
QY 2101 gcaacagccttaactgatacctaagaatggcgttgaaggagatlaaagaacttacc 2160  
DB 3350 gcaacagccttaactgatacctaagaatggcgttgaaggagatlaaagaacttacc 3409  
QY 2161 tcttcaacaggtgagcaaatatacctatgacaaatgcatgataactaacctttggagc 2220  
DB 3410 tcttcaacaggtgagcaaatatacctatgacaaatgcatgataactaacctttggagc 3469  
QY 2221 aaatagcccttaagcaatcctaataatgattccgagcaggtatgaataatgagcaacc 2280  
DB 3470 aaatagcccttaagcaatcctaataatgattccgagcaggtatgaataatgagcaacc 3529  
QY 2281 gcaacagcagtgatgatacctaagaaggcgttccatcagaaaggcgggagaagaatg 2340

DB 3530 GCACAGCAGTTGTTATGACCTCACAGGCGTTCCATTCATGCACAGCGGGAAGAAATG 3589  
QY 2341 cttcgtanaaaaggcgagcaagaaatgattataagcgggagtgatgctcaatgatt 2400  
DB 3590 cttcgtanaaaaggcgagcaagaaatgattataagcgggagtgatgctcaatgatt 3649  
QY 2401 gatcgaagcaagaagctcaataaccagatgcttcaactatataagcgggactaaccc 2460  
DB 3650 gatcgaagcaagaagctcaataaccagatgcttcaactatataagcgggactaaccc 3709  
QY 2461 ctgcttctgatacccaagccttcgcatacgaacagcaactaagaatcaatgacaccc 2520  
DB 3710 ctgcttctgatacccaagccttcgcatacgaacagcaactaagaatcaatgacaccc 3769  
QY 2521 caattcctaataatgctcgaagaacagagtgctcgaatgaatgaatgaatgaatga 2580  
DB 3770 caattcctaataatgctcgaagaacagagtgctcgaatgaatgaatgaatgaatga 3829  
QY 2581 gacaaatggggaataatcattgttataaaccacaataaactgtagcaacacat 2640  
DB 3830 gacaaatggggaataatcattgttataaaccacaataaactgtagcaacacat 3889  
QY 2641 ttgcgagcgggaaatgggcaatcgaatgacagcggtaaggtaaggaaatccaccc 2700  
DB 3890 ttgcgagcgggaaatgggcaatcgaatgacagcggtaaggtaaggaaatccaccc 3949  
QY 2701 gtlcaagcagaggaagtgctcgaagtaacaggtatcctatgataatgcttcaagag 2760  
DB 3950 gtlcaagcagaggaagtgctcgaagtaacaggtatcctatgataatgcttcaagag 4009  
QY 2761 gtaagccagaccagcgtlaaaagtaataagaaa 2794  
DB 4010 gtaagccagaccagcgtlaaaagtaataagaaa 4043

RESULT 6  
AR044549  
LOCUS AR044549 4464 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 9 from patent US 5817498.  
ACCESSION AR044549  
VERSION AR044549.1 GI:5966014  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Deweer,P. and Amory,A.  
TITLE Pullulanase Producing microorganisms  
JOURNAL Patent: US 5817498-A 9 Oct-1998;  
FEATURES  
source 1..4464  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN

Query Match 99.7%; Score 2786.4; DB 9; Length 4464;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 2793: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggaaacagcaacagatcattgctcaactatttgccttgctgattcaacct 60  
DB 1250 gatggaaacagcaacagatcattgctcaactatttgccttgctgattcaacct 1309  
QY 61 tggagctatgagatgctgccaagaagcggaggtgaggctgaatlaagatcacaacg 120  
DB 1310 tggagctatgagatgctgccaagaagcggaggtgaggctgaatlaagatcacaacg 1369  
QY 121 gctgactccttggagctgttgcaagtgctgatatcagggaaacccaagtcagga 180  
DB 1370 gctgactccttggagctgttgcaagtgctgatatcagggaaacccaagtcagga 1429

181 attatgctgcgactcgaagatgagccaaagatgtagcgctgacgctacatgatctta 240  
|||||  
Db 1430 ATTATGCTTCCACTCAAGATGGAAGATGTGAGCGGTGACCGCTACATGTGATTTA 1489  
241 agcaaaagaaatgagtgctgagctctgtagaagaaagccaaatcttctaatagaaaa 300  
|||||  
Db 1490 AGCAAAAGAAATGAGTGCTGTGGCTTGTAGAAAGAAACGCAATTTTTTATATGAAAA 1549  
301 gctgctgagagatgtagcttaaaacccgctgtaagaaagcttaattagatgcttcaacag 360  
|||||  
Db 1550 GATGCTGAGGATGACGCTAAACCCGCTTAAGCAACCTTATTAGATGCTTCAAAACGAC 1609  
361 gtgctggttaaaacttagcagcgcttaacctctgggggaagmmaaaacggcttaacgctc 420  
|||||  
Db 1610 GTGCTGCTTAACCTTAACCCGCTTAACCTTGGGAGAGNNNAACCGGCTTTACCGT 1669  
421 catgacgacacagcaaaataagatattccagtgacatctgtgaaagatgcaagcttgct 480  
|||||  
Db 1670 CATGACGACACAGCAAAATAAGGATATTCAAGTGACATCTGTGAAGATGCAAGCTTGCT 1729  
481 caagaatgaacgctgcttctggagatgacctccaacatcttggaggtccgattgg 540  
|||||  
Db 1730 CAAGATGTAACCGCTGTTTGGAGGTACCTTCCACATATTTTGGAGGCTCCGATTTGG 1789  
541 gcacctgataatcagctacttlaataaaaggtagaataaactctctaaatcttca 600  
|||||  
Db 1790 GCACCTATATATCACTACTTTATTTAAAAAGGTGACTAACATATCTCTATCAATTC 1849  
601 ggaagatctccctgaaagaaataccaaataaagtgagctttaaatagaatactgaaataat 660  
|||||  
Db 1850 GGAGATCTCTCTGAAAGAACTACCAATATTAAGTGGCTTTAATGADAGTGGAATTAAT 1909  
661 cggagttaccacatctgacacatlaattaaacgtccctcgcgaggtgcaacgctaac 720  
|||||  
Db 1910 CCGAGTTACCACTGTGACAACTTAATTAACACTCCCTCCGCGGTGCACACGCTACT 1969  
721 ttttcgtatattccgtccactcatgctgtagaagaaataaataaactctaaatcgcgat 780  
|||||  
Db 1970 TTTTCTGATATTCGCTGCACATCAAGCTATATGACAAATTAATTAATTCCTAATGCCGAT 2029  
781 ttacaagtagaaagcggggtttaaagcgtatcgtgagcggcttactctaaagggaaatacca 840  
|||||  
Db 2030 TTACAAATGAAGAAAGCGGGGTTAAACGATCTCGTAGCGGTACTTAGGGGAAGATCCA 2089  
841 gatgtgagccatctctgctcaatccaaacagatgagctcaagcaaaagcaggtgatacc 900  
|||||  
Db 2090 GATGTGAGCCATACTCTGTCCATTCAAAAGATGCTATCAGGCAAAAGCAAGTCAATACT 2149  
901 cgtaatgtgcttaattatcacagttacttaactcagagagatgacttgggaataccat 960  
|||||  
Db 2150 GGTATATGCTTAATTAATCACTACTACTATTTCAGGAGATGATCTTGGCAATACCTAT 2209  
961 acacagaagacaaacaccccttaaaagctgagcaacaaactctaaactaaagttaattgctct 1020  
|||||  
Db 2210 ACACAGAAACACACACCTTTAAAGTCTGGGACACAACTTCTACTCAAGTAAAGTCTT 2269  
1021 ctctatgacgctgaaacgggttctctgtaaaaaaaatcgtaacctatgacgagctcgccat 1080  
|||||  
Db 2270 CTTTATACAGTGCACGGCTTCTGTAACAAAATCTTACTTAAGCGGATCGGCCAT 2329  
1081 gttgtgtggaagcaacggtttaaataaaccttgaaaaaatctgtatattacatgtaagagta 1140  
|||||  
Db 2330 GGTGTGTGGAAGCAACGGCTTAATCAAAACCTTGAATAATGTAATACATGTATGACGTA 2389  
1141 acagggcaaggtctacacccgaaacggctgtgataccttaatactgcaactgtaacaaat 1200  
|||||  
Db 2390 ACAGGCCAAGGCTTACCCGACGGCTGTGATCCTTATGCAACTGGGATGCAACCAAT 2449  
1201 ggaacggagagctgattgtgtgagcctggctaaaaacagataccctgctggctggaagatgat 1260  
|||||  
Db 2450 GGAACGAGAGCATGATTGTGACCTGCAAAACACATCTCGTGGCTCGAAGAGATGAT 2509  
1261 aaacatatcagccaaagaaatataagagatgaggtcattatgaaatgagatgtgtccgagac 1320  
|||||

---

2510 AAACATATTACGCCAAGAAATATGAAAGATGAGTCACTCTTGAAATGAGATGCTCCGAC 2569  
1321 tttccattgacccctaatctgggtatgaaaaataaggaagaaatcttggcctctacagaa 1380  
|||||  
Db 2570 TTTTCATTTGACCTTAATTCGGGTATGAAAAATAAGGAAGTATTGGCTTACAGAA 2629  
1381 aaaggaacaaagggccctgacaacgtaaaagcgggagatagatctccttaaaaaaaccttgg 1440  
|||||  
Db 2630 AAAGAAACAAAGGCGCCCTGACAAAGTAAAGCGGGGATAGATTCTTAAAAACAATTGGG 2689  
1441 attactcatgttcagctatgacctgttcttgcatctaaacagtgctcgatgaaacatgacca 1500  
|||||  
Db 2690 ATTACTATGTTGAGCTTAATGCTGTTTCCATCTTAACAGCTGCTGATGAAACGATCCA 2749  
1501 acccaagataatctgggtgtaagaccctcgcaacatgatagtctctgaaaggcagatgct 1560  
|||||  
Db 2750 ACCCAATATTTGGGGTATGACCCCTGCAACTATGATGTTCTTGAAAGGCAATATGCT 2809  
1561 acaaaatcgaaatggtlaatgctcgtataaaagatllaagaaatgattcttcaatccat 1620  
|||||  
Db 2810 ACAAATCGAATGCTAATGCTGATTAAGAGATTAAGAAATGATGTTCTTCACTCCAT 2869  
1621 cgtgaacacatltggggttaacatgagatgtgtatataatcaatcccttgcacgcaaatc 1680  
|||||  
Db 2870 GGTAAACATTTGGGGTAAACATGATGTTGCTATTAATCTATACCTTTGCCACGCAATTC 2929  
1681 tctgactctgataaaatctgtaaccagaatatataacgtgaagatgataccagtaataata 1740  
|||||  
Db 2930 TCTGACTTGCATTAATTTGTACAGAAATATTATACCTGACATGATGACAGTAAATATA 2989  
1741 ccaagagatcaggtacttgaaatgaaatltgcagcngaaagccaatggttcaaaaaatct 1800  
|||||  
Db 2990 CCAACGATCAGGTACGTCGAATGAATTAATGCAACGCAATGCAATGTTCAAAAATTT 3049  
1801 attatgattcccttaagatltgggtcaatgagatcaatcttggaggttccgtttggac 1860  
|||||  
Db 3050 ATTATGATTGCTTAAATGATGTTGGGTCAATGATATGATTAATGAGGCTTCCGTTTGGAC 3109  
1861 ttaattgctgctcttggaaaaagacagatgtccaaagctgctcctcgagatctcaatgctat 1920  
|||||  
Db 3110 TTATAGCGGCTGTGGAAGAAAGACAGATGTCAAAGCTGCTCGAGCTTCAAGCTAT 3169  
1921 aatccaggaatctgacattacatgagtgacatgaaaggttgaacctctgcaatgacagat 1980  
|||||  
Db 3170 AATCCAGGAATTCACCTTACGTTAGCGTGAAGCAATGAGGCTGCAACTGCCAGAT 3229  
1981 gatcagctcttgacaaaaagagctcaaaaaagagcaatggagatgaggtgtttaaatacaat 2040  
|||||  
Db 3230 GATAGCTTGTGACAAAGAGGCTCAAAAGGCAATGGAGATGAGGTGTTAATGACAAAT 3289  
2041 ttacgaagcgttggacgagatgctcttgaatcttccgctcagatcttctggcaaggt 2100  
|||||  
Db 3290 TTACGAAGCGGTTGAGAGGCAATGCTTGAATCTTCCGCTCAAGGTTTGGCACAGAT 3349  
2101 gcaacaggttlaactgtagcaatlaagaatggcgttgaaggagatlaataatgacttacc 2160  
|||||  
Db 3350 GCAACAGGCTTAATGATGCAATTAAGAAATGCGCTTGAAGGAGATTAATGACTTTACC 3409  
2161 tcttcacaggttgaacaaatlaataatgtaacagatgataatcaatcaaccccttggagc 2220  
|||||  
Db 3410 TCTTACCAAGGTGACAAATTAATATGTCACAAAGTATGATTAATCAACCTTTGGGAC 3469  
2221 aaaaataagccctaaagcaatccctaaatgataccgaagcagatgataaatgagatgactc 2280  
|||||  
Db 3470 AAATATACCCCTAAGCATTCCTAATGATTCGAAAGCGGATGCAATTAATGATGAATC 3529  
2281 gcacaagcagltgltatgacctcaagcgcttccatltcatgcaaggcggggaagaaatg 2340  
|||||  
Db 3530 GCACAAACAGATTGTAAGACTTCACAAAGCGGTTCCATTCAATGCAAGCGGGGAAGAAATG 3589  
2341 cttcgtanaaaaggcgaacgaacgaatgattataatgcaagcgatgaggttcaatgaaatc 2400  
|||||

Db 3500 CTTCGTANAAAAAGCGCAACGACAAATAGTTATTAATGACAGGCGATCGGTCAATGATT 3649  
OY 2401 gattgagcaggaagcacaataatccagatgtttcaactatataagcggttaaccac 2460  
Db 3650 GATTGGACAGAGAAAGCTCAATATCCACATGTTTTCACACTATATTATAGCGGCTAATCCAC 3709  
OY 2461 ctgcgtcttgatccaccagccttcgcgatgacagacagtaatlgaalcaatagcacctc 2520  
Db 3710 CTTCGTCTTGTGATWCAACCCAGCCCTCCGCAATGACGACGCTAATGAATCAATGACCACTC 3769  
OY 2521 caattcccaaatatgacagagacagacagtgagcttgatataacagatcatgltatataa 2580  
Db 3770 CAATTCTTAATATAGTCCAGAGACACAGTGGCTATGAAATTGATGATCATGTTATATAA 3829  
OY 2591 gacaaatgagggaatataatctgttgttataaaccacaataaaccagtagcaaccatcaat 2640  
Db 3830 GACAAATGGGGAATATATCTGTTGTTATATACCCTAATTAACGTAACCAACATCAT 3889  
OY 2641 ttgcgcagcgggaaatgagcaatcaatgctacagcggglaaggtagagaaatccaccctt 2700  
Db 3890 TTGCGGACCGGGAATGAGGCAATCATCTACGACGCTAAGCTAGAGAAATCCACCCCTT 3949  
OY 2701 ggtcaagaagaggaggagtgctcaagttaccagttatataatgatatctctataaag 2760  
Db 3950 GGTCAAGCAGAGGAGAGTGTCCAACTACCGATATATCTATGATGATCTTATCATAGAG 4009  
OY 2761 gtaagccagacacacggtcaaaaagaatagaanaa 2794  
Db 4010 GTAACCCACGACGTAATAAAGTAATAGAAAA 4043

RESULT 7  
LOCUS AR098239 4464 bp DNA PAT 14-FEB-2001  
DEFINITION Sequence 8 from patent US 6074854.  
ACCESSION AR098239  
VERSION AR098239.1 GI:12807496  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Dwyer, P. and Amory, A.  
TITLE Pullulanase, microorganisms which produce it, processes for the  
preparation of this pullulanase and the uses thereof  
JOURNAL Patent: US 6074854-A 8 13-JUN-2000;  
FEATURES  
source location/Qualifiers  
1..4464 /organism="unknown"  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN

Query Match 99.7%; Score 2786.4; DB 9; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatgggaacagacagatcatgtgccactatcttgcctcgtggtgatatcaacct 60  
Db 1250 GATGGGAACAGACAGATCATGTGTCCACTATTTTGGCCCTGCTGATATATCAACCT 1309  
OY 61 tggagctatagatgtagccaaagaacgaggtgaggtgaaatagatccaaccg 120  
Db 1310 TGGACTATATGATGTGGCCAAAGACGAGGTGGGCTGATACGATTTCAATCAACCG 1369  
OY 121 gctgactctttggagcgtgttgcaagtgctgatalatccaggaacccaagtcagttaga 180  
Db 1370 GCTGACTCTTTTGGAGCTGTTCAGGTGCTGATATTCAGGAACCCAGTAGGTAGGA 1429  
OY 181 attacgtttcagcactcaagattggaacaaagatgtgagcgtgacgcgtacataagatta 240  
Db 1430 ATTATCGTTGCGACTCAAGATTGGACCAAAAGTGTGAGGCGCTGACCGCTACATAGATTTA 1489

OY 241 agcaaaagaaatagagtgctgtcgtctgtagaaggaaacagccaaatttltatatagaanaa 300  
Db 1490 ACCAAAGAAATAGAGGTGTGGCTTGTAGAAAGAAACAGCCAAATTTTATATAGAAAA 1549  
OY 301 gatgctgagatgacagcctaaaccgctgtgaagcaagcttattagatgcttcaaacag 360  
Db 1550 GATGCTGAGAGATGACAGCTTAACCCGCTGTAGCAACGCTTATTTAGATCTTCAACACAG 1609  
OY 361 gtcgctgtaaaacttagccagcgttaactcttgagggaaggnnaagcggttttaagtt 420  
Db 1610 GTGCTGTAACTTAAGCCAGCCGTTAACTCTTGGGGAAGNNNAACCGGCTTTACGGTT 1669  
OY 421 catgacgacacagcaaatagaataluccagtgacatcctgtgaaggatgcaagtcctgtc 480  
Db 1670 CATGACGACACAGCAAAATAGAGATATTCAGTGACATCTGTGAAGATGCAAGTCTTGCT 1729  
OY 481 caagatgtaaccgctgttcttgaggttaacctccaacataatttggaggtccgattg 540  
Db 1730 CAAGATGTAAACGCTTTTGGCAGGTACCTTCCAACATATTTTGGAGGTTCCGATTTGG 1789  
OY 541 gcaactgataatcacagtaacttataaagaagtgtaacacatctctataatctca 600  
Db 1790 GCACCTGATTAATCACAGTACTTTATTAATAAAGTGACTTAACATCTTATCAATTTCTGA 1849  
OY 601 ggaagatcttcctgaaggaaactaccataataagtggtcttaaatgatatagctgaaataat 660  
Db 1850 GGAGATCTTCTGAAAGAAAGTAACCAATTAAGGTGCTTTAAATGATAGCTGTGAATTAAT 1909  
OY 661 ccgagttaccatctgacaaacattaaattaaacgtctcgtccggcggtggtgcaagtcact 720  
Db 1910 CCGAGTTACCCACTGTGCAACATTAATTTAACGTCCCTCCGCGCGCTCACACGTCCTACT 1969  
OY 721 ttcttgatattccgtccactcatgagctcatgacagatgaacaaatgaatcttaatgagat 780  
Db 1970 TTTTCTGATATTTCCGTCCACTCTGTGACGTATGACACATTAATATATCTTAATGCGGAT 2029  
OY 781 ttacaagtagaaagcgggttlaaaaacgatactcgtgaagcttacttaaggaggaagatcca 840  
Db 2030 TTACAAAGTAGAAAGCGGGGTTAAAGAGATCTGTCAGCGTTACTAGGGGAAGATGCCA 2089  
OY 841 gatgtagccataactctgtccattcaaacagatgctatccaggaagcaggtgataacct 900  
Db 2090 GATGTAGCGCAATACCTGTCCATTCMAACGATGCTATCAGCAAAAGCAGGATACCT 2149  
OY 901 cgtaatgcttaattcatcacagtaactactatctgaagagatgactcttggaatctact 960  
Db 2150 CGTAATGTCTTATTCATCACAGTACTACTATTTCAGAGATGATCTTGGGAATACCTAT 2209  
OY 961 acacagaaagcaaacacttlaaagtlctggtccacacttctactcaaglaaatgtctct 1020  
Db 2210 ACACAGAAAGCAACACTTTAAAGTCTGGGCCAACAACTTCTACTCAAGTAATGTCTT 2269  
OY 1021 cttatagcagtgcaacgggtctctgtaacaaaaatccgtactcaatgacgcatcgggcat 1080  
Db 2270 CTTTATGACAGTGCACAGGGGTCTGTACMAAAATGCTACTATGACGGCATGCGGCAT 2329  
OY 1081 ggtgtgtgggaagcaacggttlaaataaaccttgaaaatgtgatatcaatgtcttgagga 1140  
Db 2330 GGTGTGTGGGAACCAACGGTTAATCAAAACCTTGAAATTTGGATTTACATGTTAGAGTA 2389  
OY 1141 acaagccaaggtctlaaccggaacggtgtgtgaltccttatgcaactlgtcgaatlgaccaaat 1200  
Db 2390 ACAGGCGCAAGGCTCTACCCGAACGGCTGTGATCTTTATGCAACCTCGATTGCAACCAAT 2449  
OY 1201 ggaagcagaggaatgattgtgtgacccgtgtctaaacagatccgtgctggtcggaacagtgt 1260  
Db 2450 GGAAACGAGAGGATGATTTGTGACCTGCTAAACAGATCTGCTGTGCGGAACAGTGTAT 2509  
OY 1261 aaacataatgcccacaaagaatatagaagatgaaggtcaatcatgaatgtgattgtccgtgac 1320  
Db 2510 AAACATATTACGCCAAAGATATAGAAAGTGAAGTACTATGAAATGGATGTCCGTGAC 2569  
OY 1321 ttltccatgacccatcaatcgtgtatgtaaaaaataaagggaagatatttgctcttaacagaa 1380



|||||  
Db 2570 TTTTCCATTGACCTTAATTCGGGTATGAAATAAAGGAAATTTGGCTTACAGAA 2629  
OY 1381 Aaagaaacaaagggccctgaacgtaaaagcgggga tagatcccttaaaacaaactggg 1440  
Db 2630 AAGGAACAAAGGGCCCTGACAAAGAGGAGGGGATAGATTCCTTAATAAACAATTGGG 2669  
OY 1441 attactatgltcagcttaagcttglttcgcatacaagtgltcgaatgaactgataca 1500  
Db 2690 ATTACTATGCTTACAGCTTAATCCCTGTTTCGACATCAACAGTCGATGAAACGATCCA 2749  
OY 1501 accaagaataatggggttgaaccctgcgaactaagatgltccctgaaggcag laagct 1560  
Db 2750 ACCCAATATTAATGGGGTTATGACCCCTGCAACATATGATGTTCTTGAAGGGCAATAGCT 2809  
OY 1561 acaaatgcgaatgttaagtctgataaaagagtttaaggaaatggtctcttcacacat 1620  
Db 2810 ACAATGCGAATGATGATGCTCGTATTAAGAGAGTTTAAGAAATGGTTCTTTCATCAT 2869  
OY 1621 cgtgaacacatlggggttaacaagagtgltgltcctaatacaactcttgccacgaalc 1680  
Db 2870 CGTCAACACATTCGGGGTTAAACATGATGATGCTCTATATCATACCTTTGCCACCAATC 2929  
OY 1681 tctgactctgaataatltgaaccgaataltataccgtaagatgcagagtaataata 1740  
Db 2930 TCTGACTTCGATTAATTTGACAGAAATATTAATACCGTACGATGATGACAGTATTAATA 2989  
OY 1741 ccaacggatcaggtactcgtgaataatgaatctgaacgaaagcgaatggtctcaaaaatc 1800  
Db 2990 CCAACGGATCAGGTACTGTAATGAAATGCAANGCNGAAGGCCAATGGTTCAAAAATTT 3049  
OY 1801 attatgaltcccttaagatltgggtcaaaagagtaataatgaacggtctccgttctggac 1860  
Db 3050 ATTATTAATTCCTTAAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 3109  
OY 1861 ttaatggcgtgcttgaagaaagacagatgtlccaaagctccctcgagagcttcaatgatac 1920  
Db 3110 TTAATGGCGCTGCTTGGAAAAAGACAGATGTCCAAACCTGCTGAGCTTCATGCTATT 3169  
OY 1921 aatcaggaatltgcaacttcaagtgagcaatgagcgggtggaacctctgcaatgcaat 1980  
Db 3170 AATCAGGAATTTGCACTTTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTAC 3229  
OY 1981 gatcagcttctgaagaaagagcctcaaaagagcagtgagagtaagagtglttaagacaat 2040  
Db 3230 GATCAGCTTCTGACAAAGAGAGCTCAAAAGCCATGGAGTACGGGTTTAAAGCAAT 3289  
OY 2041 ttaagaaacggttgaacggaatgtlcttgaatctccctgcacaggtcttgcagacaggt 2100  
Db 3290 TTAAGAAACGCTTGGAGAGCAATGTCTTGTATTCTTCCCTCAAGGTTTGGCAAGAT 3349  
OY 2101 gcaacaggtcttaactgaatgaatgaatgagcgttgaaggagatataatgaacttacc 2160  
Db 3350 GCAACAGGCTTAATGATGCAATTAAGATGAGCTTGAAGGAGATTAATGACTTACC 3409  
OY 2161 tcttccaaagtgagacaatlaactatgtcaaaagltcaataactacaccccttggagc 2220  
Db 3410 TCTTACACAGCTGACAAATTAATTAATCTCAAGTCAATGATTAATCACTACCCCTTTGGAC 3469  
OY 2221 aaaaataccctaaagcaatctcaatgaatctcgaagcggatgaataaagatgaactc 2280  
Db 3470 AAAATACCTTAAGCATTCCTTAATGATTCGAAGCGGATGAGATTAATAATGATGAATC 3529  
OY 2281 gcaacagagtgltgaactcacaagcgtlccatltcatgcaagcggggaagaaatg 2340  
Db 3530 GCACAAACAGTTGTTATGACCTCAAGGCGTTTCATTATGCAAGCGGGGAAGAAATG 3589  
OY 2341 ctccgttaaaaagcggcgaacgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2400  
Db 3590 CTTTCGTAATAAAGCGCCGCAACGACAAATAGTTATATCAGCGCGTGGCGTAATGAGCTTT 3649  
OY 2401 gatggagcaggaagctcaataccagatglttcaactatlaacgggtcaatccac 2460  
|||||

Db 3650 GATTGAGACAGAAAGCTCAATATCAGATGTTTCAACTATTATACGGGCTAATCCAC 3709  
OY 2461 ctctgcttgaatcacaccagccttcgcagatgaacgaatgaatgaatgaatgaatgaatga 2520  
Db 3710 CTTCGCTTGAATCAACCGAGCTTCCGATGACGACAGCTAATGAATCAATAGCACCTTC 3769  
OY 2521 caatcttaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2580  
Db 3770 CAATTCCTTAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3829  
OY 2581 gacaaatggggaataatcatgtgttataaaccaaataaataatgaatgaatgaatgaatga 2640  
Db 3830 GACAAATGGGAAATATCATTTGTTATTAACCAATAAATAAATGTAAGCAACATCAAT 3889  
OY 2641 ttgcagagcgggaatggcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2700  
Db 3890 TTGCCAGCGGGAATGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATG 3949  
OY 2701 ggtcaagcagaggaatgltcgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2760  
Db 3950 GGTCAACAGAGGAGGAGTGTCCAGTACAGGATATATGATGATGATGATGATGATGATG 4009  
OY 2761 gtaagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2794  
Db 4010 GTAAGCCAGACAGAGGATTAATAAAGTAAATAGAAA 4043  
|||||

RESULT 8  
AR098240  
LOCUS AR098240 4464 bp DNA PAT 14-FEB-2001  
DEFINITION Sequence 9 from patent US 6074854.  
ACCESSION AR098240  
VERSION AR098240.1 GI:12807497  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Deweer, P. and Amory, A.  
TITLE Pullulanase, microorganisms which produce it, processes for the preparation of this pullulanase and the uses thereof  
JOURNAL Patent: US 6074854-A 9 13-JUN-2000;  
FEATURES  
source 1..4464  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN  
Query Match 99.78; Score 2786.4; DB 9; Length 4464;  
Best Local Similarly 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Caps 0;  
OY 1 gatggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 60  
Db 1250 GATGGAGACAGAGAGATATGTCACATTTTGGCTGAGGATTAACAACCT 1309  
OY 61 tgaagctatagatgtggcacaagagcagaggtggggtcgtgatacgaatltcaatcaacccg 120  
Db 1310 TGAAGTGTATGAGATGTGGCAAAAGAGAGAGGTGGGCTGAATAGATTAATCAACCG 1369  
OY 121 gctgactccttggagctgtgtgcaagtgtgataatccagaagaaaccaagtcagtagga 180  
Db 1370 GCTGACCTCTTTGGAGCTGTTGGCAAGTCTGATATTCAGAGAAACCCAAAGTCAGTGA 1429  
OY 181 attatgcttgcacataagatltgacaaagatgltggcgctgacgctacatgataatga 240  
Db 1430 ATTATGCTTGCACTCAAGATGAGACCAAAAGATGTGAGCGCTGACCCCTACATGATTTA 1489  
OY 241 agcaaaagaaatgagtgctgtgctgtgaagaaagcaacgaatlttataatgaabaa 300  
Db 1490 AGCAAAAGAAATGAGGTGTGCTGTGAAGAAACAGCCAAATTTTATATGAAAAA 1549  
|||||



QY 301 gatctgagatgacgtacaaccgctgtaagcaacgcttattgagtcttcaaacgag 360  
Db 1550 GATGCTGAGGATGACGTAAACCCCGCTGTAGCAACGGCTTATTGTATGCTTCAACACG 1609  
QY 361 gtgcgtgtaacttagcgaaccgttaactcttgaggaaagmnaagcggtcttaacgt 420  
Db 1610 GTGCTGGTAAACCTTAGCAGCGGTTAACTCTTGGGAAAGGNNNAAGCGGCTTTACGGTT 1669  
QY 421 catgacgacagacaataaagatalatccagtgacatcgtgtaagaatgacagcttgct 480  
Db 1670 CATGACGACACAGCAATAAAGATATTCAGTACATCTGTGAAGATGCAAGTCTTGCT 1729  
QY 481 caagatgtaaccgctgcttgaggagtaacctcaacatattttggaggtccgattgg 540  
Db 1730 CAAGATGTAAACCGCTGTTTGGCAGGTACTTCAACATATTTTGGAGGTTCCGATTGG 1789  
QY 541 gcacctgtaatcacagtaacttattaaaaaagtgactcaaacatctcaatcttca 600  
Db 1790 GCACCTGATTAATCACAGTACTTATTAAAAAGTGACTTAACAAATCTCTATCAATTCCTCA 1849  
QY 601 ggaagatcttcctgaaagaaactaacatalaaagtggctttaaagtgatagctggaataat 660  
Db 1850 GGAAGATCTTCCGTGAAGAAACTCAATTAAGAAGGCTTTAAATGATAGCTGGAATAT 1909  
QY 661 ccgagttacccaatcgcgaacataattaaacagtcctccgagcggtgcacacgtcact 720  
Db 1910 CCGATTTACCCATCTGACAAACATTTAATTAAACAGTCCCTGGCGGGGTGCACACGCTACT 1969  
QY 721 ttctgttatctcgttcaactcatgacgtctatgacacaaataatcaatcttaatgcygat 780  
Db 1970 TTTTTCGTATATTCCTCCACATCATCATGACATGACAAATTAATTAATCTTAATGCGGAT 2029  
QY 781 ttacaagtaagaaagggggttaaagcgatcgcgtgacggttactctaaagggaagatca 840  
Db 2030 TTACAAGTAAGAAAGGGGGGTAAAAACGATCTCGTGACGCTTACTCTAGGGGAAGATCCA 2089  
QY 841 gatgtgaagccatactctgtcatctcaaacagaatggctatcaaggcaaacgagtgatacct 900  
Db 2090 GATGTGAGCCATACTCTGTCCATTCMAACAGATGGCTATCAGGCAAAACGAGTGTATCCT 2149  
QY 901 cgtatgtgcttaattcaacagtaactacttcaaggagatgactctgggaatacctat 960  
Db 2150 CGTAAATGCTTAATTCATCACAGTACTACTATCAAGGATGATCTTGGGAATACTCTAT 2209  
QY 961 aacacgaagacacacaacctllaaagtcgtgagcacaactctactcaagtaaatgctct 1020  
Db 2210 ACACGAAGAAAGCAAACTTTAAAGTCTGGGACCAACTTCTACTCAATAATATGTCTT 2269  
QY 1021 ctctatgacagtgcacaggggtctgttaacaaaatcgtactcaatgacggcatcgggcat 1080  
Db 2270 CTTTATGACAGTGCACACGGGTTCTGTAAACAAAATCGTACCTATGACGGCATCGGGCAT 2329  
QY 1081 ggtgtgtggaagcaacggttaataaaaccttgaaaatltgttatacaltgatgagta 1140  
Db 2330 GGTGTGTGGGAAGCAACGGTTATATCAAAACCTTGAATAATGGTATTACATGATGAGTA 2389  
QY 1141 acaagcccaaggtctcaaccgaacggtgtgtgattcccttgaacatgacatgcaacaat 1200  
Db 2390 ACAGGCCAAGGCTTACCGGAACGGCTGTGATCTTATGCAAACTCGAATTCACCAAAAT 2449  
QY 1201 ggaacgaagagacatgtgtgtgacctggtcaaaacagatcctgtctggtcgtgaacagtgt 1260  
Db 2450 GGAAAGAGAGGATGATGTGGAAGCTGCTTAAACAGATCTGCTGGTGGTAACAGTGTAT 2509  
QY 1261 aaacatataagcccaagaataataagaatgaggtcatctatgaaatgagatgtcgtgac 1320  
Db 2510 AAACATATTACCCCAAGAAATATAGAAAGTGAAGTATATTAAGAAATGATGTCCCTGAC 2569  
QY 1321 ttcttcattgacccaatctcggtatgaaataaaaggaagatatttggctcttacaaga 1380  
Db 2570 TTTTCATTGACCTTAATTCGGGTATGAAAAATTAAGGAATATTTTGGCTCTTACAGAA 2629  
QY 1381 aaaggaacaaaggccctgacaaagtaaacgaggtagatcttcttaaacaaacttggg 1440

Db 2630 AAAGGAACAAAGGCGCCCTGACAAACGTAAAGACGGGATGATTCCTTAAACAACTTGGG 2689  
QY 1441 attactcaatgctcaacttaatgctgtcttccgcatctcaaaagtgatgaactgata 1500  
Db 2690 ATTACTCATGTTCAAGTTATGCTGTGTTTGCATCTTAACAGTGTGCATAACTGATCCA 2749  
QY 1501 acccaagataatltgggtatgtaaccctcgcaactatgatacgttccctgaaggagatgct 1560  
Db 2750 ACCCAAGATAATTTGGGGTATATGCCCTCGCAACTATGATGTTCTTAAGGGCAGTATGCT 2809  
QY 1561 acaaatgtgaatgtgtaatgtctgtataaaaggtttaaggaatgtgtcttccatccat 1620  
Db 2810 ACAAAATGCCAAATGTAATCTGTATAAAGAGTTTAAAGAAATGATTTCTTTCACATCCAT 2869  
QY 1621 cgtgaacacatttgggtttaaactgtaatgtgtctataacatacactcttgcacgcaaatc 1680  
Db 2870 CGTGAACACATTTGGGTTATACATGATGATGTTGTTAATAATCATACCTTTGCCACGAATC 2929  
QY 1681 tctgacttcgataaactgtacacagatactatcaacgtacagatgataccaggttaata 1740  
Db 2930 TCTGACTTCGATTAATTTTACCAGAAATTTATACCGTACATGATGCAAGTAAATTA 2989  
QY 1741 ccaacgatacaggtactgtgaaatgaaatgcaangcnaaaggccaatggttcaaaaatc 1800  
Db 2990 CCAACGATACAGTACTCGAAATGAAATTCAGANGCNGAAAGGCCAATGTTCAAAAATTT 3049  
QY 1801 attattgattccttcaagatattgggtcaatgnglatcatattgacggtcttcgtttgac 1860  
Db 3050 ATTTATGATTCCTTAAATGATTTGGTCAATGATATCATATTCACGCTTCCTTTTGGAC 3109  
QY 1861 ttaatgagcgtcgtctgaaagaacacagatgtccaaagcgtccctcgaggtcttactat 1920  
Db 3110 TTAAATGGCGCTGTTGGAAGAACACAGATGTCCAAAGCTGCTCGGAGCTTACGTATT 3169  
QY 1921 aatcagaagatctgaccttcaagtgagacatggaaggttggaaactctgcaatgccaagat 1980  
Db 3170 AATCCAGGAATTTGACATTTACGGTGAAGCATGACGGGTGAACCTCTGACAGCCAGAT 3229  
QY 1981 gattcaactctgaacaaagaagactcaaaaaggaatggaggttgggtttttaaagtaaat 2040  
Db 3230 GATCACCTTCTGACAAAGAGGCTCAAAAAGCATGGAGTACGGGTGTTTAATGCAAT 3289  
QY 2041 ttacgaacagcgtltggaagcgcaatgtcttgaattcttccgctcaagagtttgcgaacagt 2100  
Db 3290 TTACGAACAGCGTGTGACGGCAATGTCTTGAATCTTCGCTCAAGTGTTCGACAGGT 3349  
QY 2101 gcaacaggtctaacatgacatlaagaatgagcttgaagggtgaatlaatgaccttacc 2160  
Db 3350 GCMAACGGCTTAACTGATGCAATTAAGAAATGCGGTTGAGCGGAGTATTAAATGACTTTACC 3409  
QY 2161 tcttcaacaggtgagacaatlaactatgtaacagatcagatlaactcaacccttggagc 2220  
Db 3410 TCTTCAACAGGTGACCAATTAATCTATGTCACAAAGTATGATTAACCTCTTTGGGAC 3469  
QY 2221 aaatatgacctaaagcaatcctaatagtatccgaagcgatgcgataaaatgataaccc 2280  
Db 3470 AAAATAGCCCTTAAGCAATCTTAATGATTCGAAGCGGATCGAATTAATGATGAATC 3529  
QY 2281 gcaacagcagttgttatagacctcaaaagcggttcaatcaatgcaagcggtggaagaatg 2340  
Db 3530 GCACAAGCAGTTGTTATGACCTCACAAAGCGTTCCATTATGCAAGCGGGAAGAAATG 3589  
QY 2341 ctctgtanaaaaggcgcaacgaacaaatgataatgacggagatgacgttcaatgtgctt 2400  
Db 3590 CTTTGTANAAGGCGGCAACGCAATTAATTAAGTCAAGCCATCGGCTCAATGTGATTT 3649  
QY 2401 gatitgagcaggaagcctaataatccagatgtttcaactatatagcgggtcaacccac 2460  
Db 3650 GATITGAGCAGGAACCTCAATATCCAGATGTTTCAACTATATTATGCGGGCTAATCCAC 3709  
QY 2461 ctctgcttgataccacgacctcgcatagacagcaatgaaatcaatgacacactc 2520



```

Db 2690 ATTACTCATGTTACACTTATGCTTTTCGATCTACACAGTGTGCATGATAAACTGATCCA 2749
Qy 1501 acccaagaataattggggtatagaccctcgcaactatgatgtcttgaaggagatagtc 1560
Db 2750 ACCCAAGATAATTTGGGGTATATGACCTCGCAACTATGATGCTTCGAAGGAGATATGCT 2809
Qy 1561 acaaatggaatagtgtaagtcgtatataaagagtttaaggaatgtgtcttcaactccat 1620
Db 2810 ACMAATGGCAATGTAATGCTGTATMAAAGATTAAAGGAATGTTTTCACACCAT 2869
Qy 1621 cgtgacacattggggttaacatagatgtgtctataacatacctttgcccgcgaatc 1680
Db 2870 CGTGAACACATTGGGGTATACATGATGTTGCTATTAATCATACCTTTCCACGCAATC 2929
Qy 1681 tctgacttcgaataaatttgtaccagaataatlaaccgtagatgataccaglaattata 1740
Db 2930 TCTGACTTCGATAAATTTGTACAGATAATTTATACCGTACGATGATGAGGTAATTATA 2989
Qy 1741 ccaacggaatcagtgatcctgggaatgaatctgacngcnaagagccaaatgttcaaaaatt 1800
Db 2990 CCMAAGGATCAGGTAAGTGAATGAAATTCANGCNAGAAAGCCAAATGGTTCAAAAATTT 3049
Qy 1801 attatgacttccttaagatgagtgagatgagatcatatgagagcttcggttttgcac 1860
Db 3050 ATTATTTGATTCCTTAACTATGCTGATGATGATGATGATGATGATGATGATGATGAT 3109
Qy 1861 ttaatgagcgtctgttgaataaagacagatgctcgaagcttcgagcttcatgctatt 1920
Db 3110 TTAATGGCGCTGCTTGAAAAAGACAGATGTCGCAAGGCTCGGAGGTTCAATGCTATT 3169
Qy 1921 aatcagaagaattgcaacttaagtgagtgacagatgagagcttcgagcttcgagct 1980
Db 3170 AATTCAGGATTTGCACTTTACGCTGAGCATGAGCGGATGGAACCTCTCACGCGCAGAT 3229
Qy 1981 gatacagcttcgacaaaagagctcaaaaagcaltgagagatgagagtgatgagtaataat 2040
Db 3230 GATTCAGCTTCTACAAAAGGACTCAAAAAGCATGAGGATGAGGATGATGATGATGAT 3289
Qy 2041 ttaacgaacagcgttgaacggaatgtcttcttctcgcctcaaaagtttctgcgaacgt 2100
Db 3290 TTACGAACAGCGTGTGAGACGCAATGCTTTGATTTCTCCGTCGAAGGTTTGGACAGGT 3349
Qy 2101 gcaacagagcttaactgagatgaatgaagagctgtgagagagatgataagcttacc 2160
Db 3350 GCMAAGGCTTAACTGATGATTAAGATGAGGCTTGAAGGAGATGATGATGATGATGAT 3409
Qy 2161 tcttcaacagagtgagacaaatlaactatgtcacaagtcagatgaactaaccccttggagc 2220
Db 3410 TCTTCAACAGGTGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3469
Qy 2221 aaatatagccttaagcaatlaactatgaltccgaagcgaatcgaatlaaaatgatac 2280
Db 3470 AAAATAGCCCTTAAGCAATCTTAATGATTCGAAGCGGATCGGATTAATTAATGATGAAC 3529
Qy 2281 gcaacagagctgttataagcttcaacagagcttccatcaatgaagcgggggaagaatg 2340
Db 3530 GCACAGAGCGTGTATGAGCTTACAAAGCGCTTCCATTCATGACGAGCGGGAAGAAAG 3589
Qy 2341 cctcttaaaaaagcggcgaacagaaatgattataatgacggagatgcgttcaatgagtt 2400
Db 3590 CTTTCTTAATAAAGCGGCAACACATTAATTAATGAGGCGCATTCGGGCAATGAGTTT 3649
Qy 2401 gatctgagaggaagaagctlaactatcagaatgttcaactatlatagcgggttaactccac 2460
Db 3650 GATGAGAGAGGAAAGCTCAATATCAGATGTTTCACTATTATATGCGGGGTAATTCAC 3709
Qy 2461 cctgcttcttgatcaccacaccccttcgcagatgagagacagcctaaatgaatagcacc 2520
Db 3710 GTTGGCTTTGATCACCACCTTCGCAATGAGACAGCTAATGAATCAATACCCACCTC 3769
Qy 2521 caattcctaataatgacagaagaacagtgagcctatgaaatlaactatgataatgataa 2580
Db 3770 CAATTCTTAATAATGATGACAGAAACAGAGTGGCTATGCAATTAATGATGATTAATAA 3829

```

```

Qy 2581 gacaaatggggaataatcatgttgttataaaccacaaataaactgttaagcaacatcaat 2640
Db 3830 GACAAATGGGGAATATATCATTTGTTTAAACCAATTAACCTGTAGCAACATCAAT 3889
Qy 2641 ttgagcagaggaagaatgggcaatcaatgtctagagcagtgaaagtgagagaatccacct 2700
Db 3890 TTGCGCAGCGGGAATAGGCAATCAATGCTACAGCGGTAAGCTAGAGAAATCCACCTT 3949
Qy 2701 gttcaagcagaggaagtgctccaaagtaacagatatactatgataatgtcttcaagaag 2760
Db 3950 GGTCAAGCAGAGGGAAGGTGTCCAAAGTACCAAGTATATCTATGATGATCTTCAACAG 4009
Qy 2761 gtaagccagacacagtgtaaaaagtaataaga 2794
Db 4010 GTAAAGCCAGACACGCTAAAGTAATGAAAA 4043

RESULT 10
LOCUS 189312 4464 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 9 from patent US 5721127.
ACCESSION 189312
VERSION 189312.1 GI:3409252
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4464)
AUTHORS Deweer, P. and Amory, A.
TITLE Pullulanase
JOURNAL Patent: US 5721127-A 9 24-FEB-1998;
FEATURES
Source 1..4464
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others
ORIGIN

Query Match 99.7%; Score 2786.4; DB 10; Length 4464;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatgggaacagcaacagatcattgtccactatlttgcctgtgtgtatcaact 60
Db 1250 GATGGGAACAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309
Qy 61 ttgagctctatgagatgtgcccacaaagacgaggttgggctgaaatgaatcaacag 120
Db 1310 TTGAGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
Qy 121 gctgactccttttgagctgtgtgaagtgctgatatccagaagaaaccaaagtcagtaag 180
Db 1370 GCTGACTCTTTTGGAGCTGTTGCAAGTGTGATATTCAGGAACCAACCAAGTGAAG 1429
Qy 181 attatcgttcgacatcaagatltgacaaagatgtgagcgttgacgcgtacatagatta 240
Db 1430 ATTATGTTGCGCATCAAGATTGACCAAAAGATGTGAGCGCTGACCGCTCATAGATTTA 1489
Qy 241 agcaaaaggaatgagatgtgtgtctgtagaaggaacacagccaattcttataatga 300
Db 1490 AGCAAAAGGAATGAGTGTGCTTTGTAAGAAAGCAACGCAAAATTTTATTAATGA 1549
Qy 301 gatcgtgagagatgacatcaaacgcgtgtgaagaaacgctattatagatgttcaaacag 360
Db 1550 GATGCTGAGGATCAGCTTAACCCGCTGTAAAGCAACGCTTATTAATGATGCTTAAAC 1609
Qy 361 gtgctgtttaaacttagcagccgttaactcttggggaaggnmaagcggctttagcgt 420
Db 1610 GTGCTGTAACTTAGCCAGCGGTTAACTTGGGGAAGNNNAAGCGGCTTTAGCGTT 1669
Qy 421 catgacgacacagcaataaagatattccagtgacatctgtgaagatgcaaaatcttgct 480

```



Query 2641 ttgccgagcgaggaaatggtgcaatcaatgctacagcggtlaaggtlaaggaatcaacctt 2700  
|||||  
Db 3890 TTGCCAGCGCGGAATGCGCATCATCTACGACGGGTAAAGTGAAGAACCCCTT 3949  
Query 2701 ggtcaagaagaggggaaggtgtccaaagttaccaggtatatactatgatatctctcaataag 2760  
|||||  
Db 3950 GGTCAAGAGGAGGGAAGTGTCCAAAGTACCAAGGTATATCTATGATGATATCTTATCAAGAG 4009  
Query 2761 gtaagccagacacagcgtlaaaaaagtaagaaaa 2794  
|||||  
Db 4010 GTAAGCCAGACCAACGTTAAAGTAATAGAAAA 4043  
RESULT 11  
189323 189323 4464 bp DNA PAT 10-AUG-1998  
LOCUS Sequence 8 from patent US 5721128.  
DEFINITION I89323  
ACCESSION I89323 GI:3409263  
VERSION I89323.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4464)  
AUTHORS Deweer,P. and Amory,A.  
TITLE Process for the production of novel pullulanase  
JOURNAL Patent: US 5721128-A 8 24-FEB-1998;  
FEATURES Location/Qualifiers  
source 1. 4464  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN  
Query Match 99.7%; Score 2786.4; DB 10; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Query 1 gatgggaacagacacagcatcatgtccacataatttgcctcgtgtgatacaacct 60  
|||||  
Db 1250 GATGGGAACAGACGAACCATCTTGTCCACTTTTGGCCGTGATGATATATCAACT 1309  
Query 61 tggagctactagatgtggccaaaagaacggaagtgaggctgaatacgaattcaacaaccg 120  
|||||  
Db 1310 TGGAGCTATGAGATGTGCCAAAAGACGAGGTGGGCTGAATACGATTTCAATCAACCG 1369  
Query 121 gctgactcttttgagagcgtgtgcaagtgctgatatccaagaaacccaagtgaatga 180  
|||||  
Db 1370 GCTGACTCTTTTGGAGCTGTTCAGTCTGATATTCAGGAACCCCAAGTCAAGTAA 1429  
Query 181 attacgtctgcacccaagatgtgacccaagaagtgcagcgtgacccgtacataagatta 240  
|||||  
Db 1430 ATTATCGTTCCGACTCAGATTTGGACCAAGATGTGAGCGCTGACCGCTTACATGATTTA 1489  
Query 241 agcaaaaggaatagagtgctgctgtgtagaagaaacagccaatctttataatgaanaa 300  
|||||  
Db 1490 AGCAAAAGGAATAGAGTGTGCTGTAGAAAGAAACAGCAAAATTTTATATGAAAAA 1549  
Query 301 gatgtctgagatgtgacgtcaaacccgctgtaagcaacgcttatattgattgctcaaccg 360  
|||||  
Db 1550 GATGTCTGAGGATGACGTAAACCCGCTGTAAACAGGCTTATTTAGATCTTCAACCG 1609  
Query 361 gtgctgtgtaaaacttagccagcgcgttaacctcttgagggaagannaagcggcttaacgt 420  
|||||  
Db 1610 GTGCTGTAAACTTAGCCAGCGCTTAACCTCTGGGGAAGNNNAAGCGCGCTTTACGCTT 1669  
Query 421 catgacgacacagcaataaagataatccagtgacatctgtgaaggaatgcaagtcttgct 480  
|||||  
Db 1670 CATGACGACACAGCAAAATAGATTTTCACATGTGTGAAGGATGCAAGTCTTGTGCT 1729  
Query 481 caaagtgtaaacccgctgttttgcaagtaacctccaacatattttggaggtccgattgg 540  
|||||  
Db 1730 CAAGATGTAAACCGCTGTGTTGGAGGTACCTTCAACATATTTTGGAGGTCCGATTTGG 1789

Query 541 gcaactgtataacagtaactttaaataaaggctgaactaacaatctctataacttca 600  
|||||  
Db 1790 GCACCTGATTAATCACAGTACTTTATTTAAAAAGGTGACTTAACAAATCTATCAATTTCA 1849  
Query 601 ggaagatctctcgaaggaaactcaataaagaatggtcttaataatgaatgctgaataat 660  
|||||  
Db 1850 GGAAGATCTTCCCTGAAAGGAAACCAATTAAGTGGCTTTTAAATGATAGCTGGAATTAAT 1909  
Query 661 ccgaattacccaactgacacacataaattaaagcccgccggcgagtcacagctcact 720  
|||||  
Db 1910 CCGAGTTACCACTGTGACACATTAATTAACAGTCCCGCGGCGGACACGCTCACT 1969  
Query 721 ttcttgtaattccgtccacatcgtacatgacatgacacataataatccaaatgctgag 780  
|||||  
Db 1970 TTTTGATATATTCGCTCCACCTCATGACGCTATGACACAAATTAATATCTATGCGGAT 2029  
Query 781 ttcaagtagaagaacggggttaaaaacggaatcgtgaacggttactctlaagggaagatcca 840  
|||||  
Db 2030 TTACAAAGTAGAAAGCGGGGTTAAACGGATCTGACGCTTACTCTAGGGGAAGATCCA 2089  
Query 841 gatgtgaaccataactctgtccatctcaaaagatggtctacaggaaggtgataacct 900  
|||||  
Db 2090 GATGTGAGCCATCTCTGTCCATTCAAACAGATGGCTATGAGCAAAAGCAGGTGATCCT 2149  
Query 901 cgtaaatgtctaaatcatalcacaagtaactatccagagatgactcttggaataactat 960  
|||||  
Db 2150 CGTAATGTCTTAATTCATACACAGTACTACTATTCAGAGATGATCTTGGAATACCTAT 2209  
Query 961 acacagaagaacacaacacttaaaagctcgtggcacaacacttactcaaglaaagtctct 1020  
|||||  
Db 2210 ACACAGAAAGCAACCACTTAAAGCTGTGGCACCAACTTACTCACTCAATTAATGTTCTT 2269  
Query 1021 cttatgacagtgacacaggggtctgtgaaacaaatctgacataacagcagcagccat 1080  
|||||  
Db 2270 CTTTATGACAGTGCACAGGGGTTCTGTAAACAAATGCTATGACGGCATCGGCCAT 2329  
Query 1081 ggtgtgtggaagcaacgggttaatacaaaccttgaaatctglatatcgtatgaagta 1140  
|||||  
Db 2330 GGTGTGTGGGAAGCAACGGTTAATCAAAACCTTGAAATTTGGATTAATGATGAGTA 2389  
Query 1141 acagggcaaggctctaacccgaacgctgttgatccttatgacaactgtgacacaat 1200  
|||||  
Db 2390 ACAGGCGCAAGGCTTACCCGAACGGCTGTGATCTTATGCAACTGCGATTCACCAAAAT 2449  
Query 1201 ggaacgagaagatgaattgtgacactggttaaaacagatcccgctggtcggaacagtgat 1260  
|||||  
Db 2450 GGAACGAGAGCATGATGTGTGACCTGGCTAAACAGATCTGCTGGTGAACAGTGTAT 2509  
Query 1261 aaacataatcagccaagaagataaagaatgaagtcacatcatalatgaagaatgagtgac 1320  
|||||  
Db 2510 AAACATATTACGCCCAAGAAATATAGAAATGAGGTGATCTATGAATATGATGCTCGTAC 2569  
Query 1321 ttltccattgacccttaattcgggtatgaanaaataaagggaaglatlttgctcttacaga 1380  
|||||  
Db 2570 TTTTCATATGACCCCTTAATTCGGGTATGAAAAATTAAGGGAAGTATTTGGCTCTTACGAA 2629  
Query 1381 aaagaaacaaaggccctgaacgtaacgtaaaagacgggaatgattctcttaaaacacttgg 1440  
|||||  
Db 2630 AAAGGAACAAAGGCCCTCAACACGTAAGACGGGATGATTTCTTAAACAAACTTGGG 2689  
Query 1441 attactcatgttcagctttagctgttctgcacatccaaagtgatgaglaaactgtacca 1500  
|||||  
Db 2690 ATTACTCATGTTACAGCTTATGCTGTGTTTTCGATCTTAACAGTGTGATGAACACTGATCA 2749  
Query 1501 acccaagaataatggggtatgaacccctgcgaactatgatagtctcgtgaaggcaglatgct 1560  
|||||  
Db 2750 ACCCAAGATAATTTGGGTTATGACCTTCGCAACTATGATGATTTCTTCAAGGAGTATGCT 2809  
Query 1561 acaaatgtgaatgtgaatgctcgtataaagaaggttaagaagaatgtttttcacaccat 1620  
|||||  
Db 2810 ACAAAATGCAATGTATGCTCGTATTAAGAGTTTAAGGAATGATTTCTTACATCTCAT 2869

1621 cgtgaacacatttggttaacatgatgttgtatataataacacatttgcacagcaatc 1680  
|||||  
Db 2670 CGTGAACACATTGGTGAACATGATGTGTCTATATATACCTTTGCCACCAATTC 2929  
1681 tctgaactcgaataaattgtaccgaatataataacacatgcacagcaatc 1740  
|||||  
Db 2930 TCTGACTTCGATTAATAATTGATACGAAATTTATACGATGATGACGATTAATA 2999  
1741 ccaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800  
|||||  
Db 2990 CCAACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3049  
1801 atcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860  
|||||  
Db 3050 ATTATTTATTTCTCTTAACTATTTGGTCAATGATATATGACGCTTCCGTTTTCAC 3109  
1861 tcaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1920  
|||||  
Db 3110 TTATGAGCGGCTTGGTGAAGACAGATGTCGAACGCTGCTGGAAGCTTCAATGAT 3169  
1921 atcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980  
|||||  
Db 3170 ATTCAGGAATTCACCTTACGCTGACGACATGACGCTGACGCTGACGCTGACGAT 3229  
1981 gatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040  
|||||  
Db 3230 GATCAGCTTCTGACAAAGAGCTCAAAAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAG 3289  
2041 tcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100  
|||||  
Db 3290 TTACGAAGCGGCTTGGAGAGGATGTTGATGCTTCCGCTCAAGTTTTCACAGAGT 3349  
2101 gcaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2160  
|||||  
Db 3350 GCAACAGCGCTTAACCTGATGCAATTAAGCAATGCTGACGCTGACGCTGACGCTGAC 3409  
2161 tctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220  
|||||  
Db 3410 TCTTACAGAGCTGACAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 3469  
2221 aaaaac 2280  
|||||  
Db 3470 AAAAAACCCCTTAAGCATTCCTAATGATTCGGAAGCGATGGAATTAATGATGATGATG 3529  
2281 gcaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340  
|||||  
Db 3530 GCACAAACAGCTTTATGACCTCACAAGCGCTTCCATTCATTCATTCATTCATTCATTC 3589  
2341 ctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400  
|||||  
Db 3590 CTTCCTTAAAAAGCGCGCACACGCAATGATGATGATGATGATGATGATGATGATGATG 3649  
2401 gatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2460  
|||||  
Db 3650 GATGAGAGCAGGAAGCTCAATATCCAGATGTTTCAATATATATACGGGCTAATCCAC 3709  
2461 ctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2520  
|||||  
Db 3710 CTTCCTTAAAAAGCGCGCACACGCAATGATGATGATGATGATGATGATGATGATGATG 3769  
2521 caatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2580  
|||||  
Db 3770 CAATTCCTTAAAAAGCGCGCACACGCAATGATGATGATGATGATGATGATGATGATGATG 3829  
2581 gacaaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2640  
|||||  
Db 3830 GACAAATGAGGGAATATCATTTGTTTATTAACCAATTAATTAATTAATTAATTAATTA 3889  
2641 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700  
|||||  
Db 3890 TTCCCGCAGCGGAATTCGCAATCAATGATGATGATGATGATGATGATGATGATGATGATG 3949  
2701 gacaaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2760

|||||  
Db 3950 GGTCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4009  
2761 gtaagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2794  
|||||  
Db 4010 GTAAGCCAGAGCAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4043  
RESULT 12  
189324 189324 4464 bp DNA PART 10-AUG-1998  
LOCUS  
DEFINITION Sequence 9 from patent US 5721128.  
ACCESSION 189324  
VERSION 189324.1 GI:3409264  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 4464)  
AUTHORS Deweer, P., and Amory, A.  
TITLE Process for the production of novel pullulanase  
JOURNAL Patent: US 5721128-A 9 24-FEB-1998;  
FEATURES Location/Qualifiers  
source 1..4464  
BASE COUNT 1355 a 855 c 1039 g 1187 t 28 others  
ORIGIN  
Query Match 99.7%; Score 2786.4; DB 10; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Query 1 gatggaacacgac 60  
|||||  
Db 1250 GATGGAACACGACACGATCATTTGCTACATTTTGGCTGCTGATATCAACT 1309  
61 tggagctatgagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120  
|||||  
Db 1310 TGCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1369  
121 gctgac 180  
|||||  
Db 1370 GCTGACTCTTTTGGAGCTGTTGCAAGTGTGATTTCCAGGAAACCCAGATCAGTGA 1429  
181 atcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240  
|||||  
Db 1430 ATTATGCTTGCACACTCAAGATTTGACCAAAAGATGAGCGCTGACGCTACATGATTTA 1489  
241 agcaaaagaaatgagatgagcagcagcagcagcagcagcagcagcagcagcagcagc 300  
|||||  
Db 1490 AGCAAAAGAAATGAGTGTGCTGTTGAGAAAGAAACAGCCAAATTTTATTAATGA 1549  
301 gatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360  
|||||  
Db 1550 GATGCTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1609  
361 gtcgctgtaaaactgaacagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
|||||  
Db 1610 GTGCTGTTAACTTACCAAGCGCTTAACCTTGGGGAAGNNNAAGCGGCTTTACGCTT 1669  
421 catgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
|||||  
Db 1670 CATGACACACACAGCAATTAAGATATTCAGTGACATCTGTGAAGATGCAAGCTTGG 1729  
481 caagatgaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
|||||  
Db 1730 CAAGATGTAACCGCTGTTTGGAGGTACCTTCCAAATATTTTGGAGTTCGATGG 1789  
541 gacactgataacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600  
|||||  
Db 1790 GCAAGCTATATACAGCTACTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1849

QY 601 ggaatctctctgaaggaactaccataataaagtgccttaaatgataagcttgataat 660  
Db 1850 GGAATCTCTCTGAAGGAACATACCAATATAAGGCTTTAAAGATAGCTGGATATAT 1909  
QY 661 ccgaatcccatctgacaacatlaattaaacagtcctcgcgagcgatgacaacagct 720  
Db 1910 CCGAATCCCATCTGACACAAATTAATTAACAGTCCCTGGCGGGGACACGTCCT 1969  
QY 721 ttctgataatctcgctcactcaatgagctctatgacacatlaataatcctaatgagat 780  
Db 1970 TTTTCGTATATCCGTCACATGACAGTATGACACAAATTAATTAATCCTATGCGGAT 2029  
QY 781 ttacaagtagaagaagcgaggttaaaagagatctcgtagcaggttactctgaggaagatcca 840  
Db 2030 TTACAAAGTAGAAGCGGGGTTAAACCGATCTCGAGGGTTACTCTAGGGGAATCCCA 2089  
QY 841 gatctgaaccactactctgtcatcttaacacagatgagctatcaggaagaagtgatact 900  
Db 2090 GATGTGACCCATCTCTCTCTCATTTCAACAGATGGCTATCAGGCAAGCAGGTATACCT 2149  
QY 901 cgtaatgtcttaattcaatcagagtaactaactcaagagatgatactctggaatacctat 960  
Db 2150 CGTATGTGCTTAATATCATACAGACTACTATTCAGAGATGATCTGGGAATACCTAT 2209  
QY 961 acacaagaagaacaacacttaaaagctcgagcaaccaactctactcaagtaaatgtctct 1020  
Db 2210 ACACAGAAAGCAACAACTTTAAAGCTTGCGGCCACCACTTCTACTCAAGTAATGTCTT 2269  
QY 1021 cttatagcagtgacaaggggtctgttaacaaaaactctacacataagcaggaacgggcat 1080  
Db 2270 CTTTATGACAGTGCACAGGGTCTGTGTAACAAAAATGTAACCTATACGCGCATGGGCCAT 2329  
QY 1081 ggtgtgtggaagaacaagcttaataaaaaacttgaaaaacttggtatataatgataagta 1140  
Db 2330 GGTGTGTGGAAGCAACCGTTAATCAAAAACCTTGAAAATGGTATATCATGATAGAGTA 2389  
QY 1141 acagggccaaggtctcaccgaagcggtgttgatctcttatgcaactgagatgacccaat 1200  
Db 2390 ACAGGCCAAGGCTCTACCCGGAACGGCTGTGATCCTTATGCAACGCGATTCACCAAAAT 2449  
QY 1201 ggaagaaagacatgattgtgtgacctgtgctaaacagatccctgctggtctggaacggtat 1260  
Db 2450 GGAAGCAAGCATGATTGTGACCTGTGCTTAAACAGATCCGTGCGCTGGAACAGTGAT 2509  
QY 1261 aaacatatacgaagaataataaagaatgagctatcatalgaaatgatactccgtgac 1320  
Db 2510 AAACATATTAAGCCAAATAATAGACATGAGTCACTATGAATGATGATTCGCTGAC 2569  
QY 1321 ttcccatctgacctaatctcggtgtgtaaaaaataaagggaagatattggtctctacaga 1380  
Db 2570 TTTTCCATTGACCCCTAATTCGGGTTGAAAAATAAAGGAAAGTATTTGGCTCTTACAGAA 2629  
QY 1381 aaggaagaagaagggccctgacaagaatgaagcgggagatagattccttaaaaaaacttgg 1440  
Db 2630 AAAGGAAGAAAGGGCCCTGACACAGTAAGAGGGGATAGATTCTTAAACAAACCTTGGG 2689  
QY 1441 attactcatctgcaagctatgctgttctgcacatcaacagtgctgagaaactgatacca 1500  
Db 2690 ATTACTCATGTTCAAGCTTATGGCTGTTTCGCATCTAACAGTGTGAGTAAACTGATCCA 2749  
QY 1501 acccaagaataatggtgttatagaacctcgcaactatgattgctcgtgaaggagcatagct 1560  
Db 2750 ACCCAAGTAATTTGGGGTTATGACCCCTGCAACTATGATGTTCTCGAAGGGGAGATATGT 2809  
QY 1561 acaaatggaatgtaatgctgctgataaaagatcttaaggaagaatggtcttccatccat 1620  
Db 2810 ACAAAATGGCAATGTAATGCTGCTATAAAGAGTTTAAAGGAATGTTCTTACCTCCAT 2869  
QY 1621 cgtgaacacatctgggttataacatgagatgtctgctataatcattacaccttgccacgcaaatc 1680  
Db 2870 CGTGAACACATTTGGGGTTAAACATGATGTGTCTATTAATCAATACCTTTGCCACGCAAAATC 2929  
QY 1681 tctgaactcgtataaaattgtaccagaataattacacgtacgaatgataccaggttaattata 1740

Db 2930 TCTGACTTCGATAAAATTTGTACACAAATATTATACCGTACATGATGACGGTAAATTATA 2989  
QY 1741 ccaacggatcaagtaactggaataatgacangnagaagccaatggttcaaaaatt 1800  
Db 2990 CCAACGGATCAAGCTACTGGAATGAATTTGCANCGNAGAAAGCCAAATGTTCAAAAATTT 3049  
QY 1801 altactgtatcccttaaglatatggttcaatgagatcaatattgacggtctcgtttgac 1860  
Db 3050 ATTATGATATCCCTTAAGTATGGGTCAATGAGTATCATATTGACGGGCTTCGTTTGAC 3109  
QY 1861 ttaatggcgctgtcttggaagaagaacagatgtccaaagctgtcctcgagcttcaatgtatc 1920  
Db 3110 TTAATGGCGCTGCTTGAAAAACACAGATGTCCTCAAGCTCCCTCGAGCTTATGCTAAT 3169  
QY 1921 aatccagaatctgacattcagttgagtgacatgaaagggatggaacctctcactgccaagt 1980  
Db 3170 AATCCAGAAATTTGCATTTACGGTCAACCATGACGCGGTGGAACCTCTGCACCTGCCAGAT 3229  
QY 1981 gatcagctcttgacaagaagagctcaaaaaggcatggagtagcggtgttlaatgacaat 2040  
Db 3230 GATCAGCTTCTGACAAAAGAGACTCAAAAAGGCATGGAGTAGCGGTTTATGACAAAT 3289  
QY 2041 ttaagaaacggtgtggaaggaatgtcttgaattcttcgcgtcaaggtttgagacaagt 2100  
Db 3290 TTACGAACCGCTTGACGCGCAATGCTTTGATTTCTCCGCTCAAGGTTTCCGACAGCT 3349  
QY 2101 gcaacaggtctaacgtatgataatgaatgagctgttgaagggaagatlaataacttacc 2160  
Db 3350 GCACACGCTTAACTGATGACAAATTAAGATGCGCTTGAGGGAGATTAATGACTTTTACC 3409  
QY 2161 tcttaccacagtgtagaacaatlaactatgctacaaagctacatgataactacaccttggagc 2220  
Db 3410 TCTTCACCGTGTGACAAATTAATGATGTCACAAAGCATGATTAACCTTTGGGAC 3469  
QY 2221 aaaaatgaccttaagcaatccctaatgattccgaagcgatcgatlaaaatgataagctc 2280  
Db 3470 AAAAATAGCCCTTAAGCAATCTAATGATTCGAAAGCGATCGGATTAATAATGATGAACTC 3529  
QY 2281 gacaaagcaggtgttatgacctcacaaggcgttccatlaatgacaaggcgggaagaagaatg 2340  
Db 3530 GCACAAAGCATGTTGATGACCTTCACAAAGCGCTTCCATTCATCCAAAGCGGGAAGAAATG 3589  
QY 2341 ctctgtanaaaagcggcaacgacaatgattataatgacaggtatgctcaatgattc 2400  
Db 3590 CTTGCTANAAAAGCGCGCAACAGCAATATTATTAAGCAGGATCGGCTCAATGATTT 3649  
QY 2401 gattggaagcagaagctcaatataccagaatgtttcaactatataagcggctaatccac 2460  
Db 3650 GATTGGAGCAGGAAGCTCAATATCCAGATGTTTCAACTATTAATATAGCGGCTAATCCAC 3709  
QY 2461 ctgtctgtatgacacccagccttcgcagtagaagaagctaaatgaaatcaatgccaacc 2520  
Db 3710 CTTTCGCTTGATCCACCCCTTCGCAATGACAGCAAGCTTAATGAATTAATAGCCACCTC 3769  
QY 2521 caattcctaataatgctcagaagacaacagtgacctatgaaatlaactgatactgtaataa 2580  
Db 3770 CAATTCCTTAATATGTCACAGAAACACAGTGGCTATGAATTAAGTATCATGTTAATAA 3829  
QY 2581 gacaaatggggaataatcatgtgttctataaaccacaataaaactgttagcaacatcaat 2640  
Db 3830 GACAAATGGGGAATATCATTTGTTTAAACCCAAATAAATAGTATACCAACATCAAT 3889  
QY 2641 ttgcgaagcgggaaatlgggcaatcaatgttataagagcggtlaaggaatccacacct 2700  
Db 3890 TTGCGCAGCGGGAATGGCAATCAATGCTACGACGGGTAAAGGTACGAATAATCCACCTT 3949  
QY 2701 ggtcaagcagaaggaagtgttccaagtacaggtatatacctatgatactcctacataagag 2760  
Db 3950 GGTCAAGCAGAGGAAGGTCTCAAGTACCAAGTATATGATGATCTTATCAAGAG 4009  
QY 2761 gtaagccacagccaaggttaaaaagtataagaaa 2794







D	2390	CCACGGATCAGCTACTGCAAAATGAATTCGANCNGCAGAAAGGCCAATGGTTCAAAAATTTT	304.9
Q	1801	attatgaattcccttaagtaatttggtgcataagtalcatatattgaacggtccgtttgac	1860
D	3050	ATTATTGATTCCTTAAAGTTTGGGTCATGAGTATCATATTGACGGCTTCGGTTTGAC	3109
Q	1861	ttaatgacgtctgttgcataaagacacgaltgccaaagtcctctggaagcttcattgcat	1920
D	3110	TTAATGAGCGCTGCTTGAAAAAGACAGATGTCCTCAAAAGTGCGCTCGAGCTTCATGCTATT	3169
Q	1921	aatccgaagatctgcacttaacggttgagccaatgaaaggttgaaactctgcactgcagat	1960
D	3170	AATCCAGCAATTCACCTTAAACGGTGGCCATGACGGGGGGAACCTCGCACTGGCAGAT	3229
Q	1981	gatacagctctgcacaaagagcttcaaaaagccalggggatgagtggtttaaagacat	2040
D	3230	GATCAGCTTTGACAAAAGAGCTCAAAAAGCATGGAGTATGCGGTGTTTAATGACANT	3289
Q	2041	ttaacaaacgcgtttgacagggcaatgtctcttgaattcttcgcctcaaggttttgcacagt	2100
D	3290	TTACGAAGCGTTTGACGGGCATGTGTTGATTTCTCCGCTCAAGGTTTGCGACAGT	3349
Q	2101	gcaacaggtctaacctgatgcataaagaatggcgttgagggagatlaatgaactaac	2160
D	3350	GCAACAGGCTTTAACTGATGCAATTAAGAAATGGCGTTGAGGGAGATTAATCACCTTTAC	3409
Q	2161	tccttcacaaagttagaacaattaaactctgcacaaagtcataagataaacctcttggac	2220
D	3410	TCCTTACCCAGGTGAGACATTAACCTATGTCACAGTCATGATACCTACACCTTTGGAC	3469
Q	2221	aaaaatagcccttaagcaatccctaaatgaattccgaagcggatcggatlaaaatgatatc	2280
D	3470	AAATATGCCCTTAAGCAATCCTATGATTCGCAAGCGGATCGGATTAATATGATGATAC	3529
Q	2281	gcaacaaacagtgttatagacctcaaaagcgttccatctaatgcgaagcggggaaagatg	2340
D	3530	GCAACAACTGTTTATGACCTTCACAAAGCGTTCCTTATATCAAGCGCGGCAAAAAATG	3589
Q	2341	cttcgtanaaaagcgcgaacacgaatagtatataatgcagagcgatgcgttcaatgaagtt	2400
D	3590	CTTGCTANAAAAAGCGGCACACGCAATAGTTATATGACGGCGATCGGTCAATCAGTTT	3649
Q	2401	gattgcagcaggaagctcaatactccaagatgttttccaactatataagcgggttaatcac	2460
D	3650	GATTGGAGCAGGAAGCTCAATATCCAGATGTTTCACTAATTATTAAGCGGCTATCCAC	3709
Q	2461	cttcgctctgatcaaccacgcttcgcgaatgcagacagctaatgaaatcaatagccactc	2520
D	3710	CTTCGCTTATATACCCAGCGCTTCCCATGACACACGCTAATGAATCAATATGGCCACCTC	3769
Q	2521	caattctctaattagtcgaaggaacaaagtggtctatgaattaacatgcatatgtttaaata	2580
D	3770	CAATTCTTAATATGTCACAGAACACAGTGGCTATGAATTAATCATGATCTTATATAA	3829
Q	2581	gacacaaatggggaaatatcatctgttctttaaaccacaaataaaactgttagcaaccatcat	2640
D	3830	GACAAATGGGGAATATCATCTGTGTTTATTAACCCAAATATAAACTGTAGCAACATCATAT	3889
Q	2641	ttgcgaagcgggaaatgggcataatgctcaagcgttaagcttaggagaatccaccctt	2700
D	3890	TTTCCGAGCGGGAATATGCGCAATCAATGTACACGCGGTAACTGTAGAGAAATCCACCTT	3949
Q	2701	ggtaacgaagggaagtgctccaagttacccaagatattctatgatatcccttcaataagag	2760
D	3950	GGTAAACCAAGAGGAATGTCCAAAGTACAGGATATATCATATGATCCTTCATCAAGAG	4009
Q	2761	gtaagccagacacacggtlaaaagtataagaaa 2794	
D	4010	GTAAGCCACAGCACGCTATAAAAGTATATGAAA 4043	
RESULT	14		
	193620		

LOCUS	193620	4464 bp	DNA	PAT	01-DEC-1998
DEFINITION	Sequence 9 from patent US 5731174.				
ACCESSION	193620				
VERSION	193620.1	GI:3938090			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 4464)				
AUTHORS	Deheer, P. and Amory, A.				
TITLE	Process for the saccharification of starch				
JOURNAL	Patent: US 5731174-A 9 24-MAR-1998;				
FEATURES	Location/Qualifiers				
source	1..4464				
BASE COUNT	1355 a	855 c	1039 g	1187 t	28 others
ORIGIN					
Query Match	99.7%	Score 2786.4;	DB 10;	Length 4464;	
Best Local Similarity	100.0%	Pred. No. 0;			
Matches 2793;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Db	1	gatggaaacagcagcaagcatcatctgtccactatcttgcgcctcgtcgtgattatcaact	60		
Db	1250	gatggaaacagcagcaagcatcatctgtccactatcttgcgcctcgtcgtgattatcaact	1309		
Db	61	tggagctctatgatatgtgtgccaaaagaacggaagtgaggctgtgaatcagattccaacccg	120		
Db	1310	tggagctctatgatatgtgtgccaaaagaacggaagtgaggctgtgaatcagattccaacccg	1369		
Db	121	gctgaactcttggagctgtgtcaagtgctgtatcttcagagaaacccaagtacagtagga	180		
Db	1370	gctgaactcttggagctgtgtcaagtgctgtatcttcagagaaacccaagtacagtagga	1429		
Db	181	attatcgtctgcactcaagaattggaaccaagaatgtgaagcgtcgtacgcgtacataagatta	240		
Db	1430	attatcgtctgcactcaagaattggaaccaagaatgtgaagcgtcgtacgcgtacataagatta	1489		
Db	241	agcaaaaggaagaatgagagtgctgtgttagaagaagaacacccaatttctataagaataa	300		
Db	1490	agcaaaaggaagaatgagagtgctgtgttagaagaagaacacccaatttctataagaataa	1549		
Db	301	gactcgtgagatgagcagctaaacccgctgtgaacaaagcttatcttgaagctccaacag	360		
Db	1550	gactcgtgagatgagcagctaaacccgctgtgaacaaagcttatcttgaagctccaacag	1609		
Db	361	gtgctgtgtaaaccttgcagcagcgtttaaactcttgggaaagmnaagcgcgttcaagtt	420		
Db	1610	gtgctgtgtaaaccttgcagcagcgtttaaactcttgggaaagmnaagcgcgttcaagtt	1669		
Db	421	catgacgacacagcaaaataagatatctcagatgaacatctgtgaagaatgcaagcttgat	480		
Db	1670	catgacgacacacagcaaaataagatatctcagatgaacatctgtgaagaatgcaagcttgat	1729		
Db	481	caagaatgataacgcgtcttctggcaggttacctccaacatatcttggagagtcgcattgg	540		
Db	1730	caagaatgataacgcgtcttctggcaggttacctccaacatatcttggagagtcgcattgg	1789		
Db	541	gcaacgtgataatcaagtaacttatttaaanaaaggctgaactaacaactctatcaattctca	600		
Db	1790	gcaacgtgataatcaagtaacttatttaaanaaaggctgaactaacaactctatcaattctca	1849		
Db	601	ggaagatcttctcgtgaaggaactacaataataagtgcttlaaatgataagctggaataat	660		
Db	1850	ggaagatcttctcgtgaaggaactacaataataagtgcttlaaatgataagctggaataat	1909		
Db	661	cgaagatcccaatctcaacacatlaattaaagatccctgcgggggggtgacaagttcaat	720		
Db	1910	cgaagatcccaatctcaacacatlaattaaagatccctgcgggggggtgacaagttcaat	1969		
Db	721	ttctcgtatattcgcgcacatcagagctatgacacaataataatcctaattgagat	780		



SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 2784)  
AUTHORS Dwever, P. and Amory, A.  
TITLE Pululanase, micro-organisms producing the same, method for preparation thereof as well as its use  
JOURNAL Patent: EP 0605040-A 10 06-JUL-1994;  
SOLVAY (BE)  
COMMENT Other publication JP 6217770 940809  
Other publication CA 2112028 940629  
Other publication CN 1090325 940803  
Other publication AU 527593 940707  
Other publication FI 935900 940629  
Other publication BE 1007723 951010  
Other publication BE 1007313 950516  
Other publication BE 1006483 940913.  
FEATURES  
Location/Qualifiers  
1..2784  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 864 a 558 c 635 g 721 t 6 others  
ORIGIN

Query Match 99.4%; Score 2776.4; DB 9; Length 2784;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2783; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatgggaacacgacacgacatcatcttccactatcttgcctcgtcgtgagatatacaact 60  
DB 1 GATGGGAACACGACACGACATCTTCCACTATTTTGGCCGCGTGATATCAACT 60  
QY 61 tggagctatgagtggtgccaagaagagtggtggtcgaatagatccaacccg 120  
DB 61 TGGAGCTATGATGTGGCCAAAAGACGAGGTGGGCTGAATAGCATTTCAACCG 120  
QY 121 gctgacatcttctgagctctgtgcaagtgctgatatccaggaacccaagtcaggtagga 180  
DB 121 GCTGACATCTTTGGAGCTGTTGCAAGTGTGATTTCCAGGAACCCAAAGTAGGA 180  
QY 181 attatcgttcgacccaagaatgtgacccaagaatgtgagcgctgcacgctacatagatta 240  
DB 181 ATTATCGTTCACACCAAGATTGGACCAAGATGTGAGCGCTGACCGCTACATGATTTA 240  
QY 241 accaagaagaatgagtgctgtgctgtgagaaggaacagccaatttttaataagaanaa 300  
DB 241 ACCAAGAAGAAATGAGGTGTGCTTTACAGGAACAGCCAAATTTTATATGATAAAA 300  
QY 301 gatgctgagatgacgctaaacccgctgtgaacgaagcttattagatgctcaaacag 360  
DB 301 GATGCTGAGATGACGCTAAACCCGCTGTAAACAAGCTTATTTGATGCTTCAAAACAG 360  
QY 361 gtgctgttaactatagcagccgttaactcttggggaaggaaggaagggccttaagctt 420  
DB 361 GTGCTGTAACTTAACTTAAAGCCGCTTAACTCTTGGGGAAGNNNAAGCCGCTTAAAGGTT 420  
QY 421 catgacgacacgacaataaagatalatccagtgacatctgtgaagatgcaagcttggt 480  
DB 421 CATGACGACACGACAATAAAGATATTCAGTGACATCTGTAAAGATGCAAGCTTGGT 480  
QY 481 caagaatgaacccgtgtttggcagagtaacctccaacatatcttggaggttcgattgg 540  
DB 481 CAAGATGAACCGGTGTTTGGCAGGTACCTTCAACATATTTTGGAGAGTTCGCAATGG 540  
QY 541 gcaactgtaatacagtaacttataaanaaggtgactaacaacatctcataacttcta 600  
DB 541 GCACCTGATTAATCAGATCTTATTAATAAAGGTGACTTAACAATCTCATATTAATTTCTA 600  
QY 601 ggaagatcttcctgaaggaatacctaataaagaatgagcttcaatgataagctggaataat 660  
DB 601 GGAGATCTTCTCGAAGGAACACTCAATATAAGTGGCTTTAAATGATAGCTGGAATTAAT 660

QY 661 ccgagttaccatctgacacaatlaattaaagctccctgcgagcggtgcacagctact 720  
DB 661 CCGAGTTACCATCTGACAAACATTAATTAACAGTCCCTGCCGGCGGTGCACACGCTACT 720  
QY 721 ttctgtaattccgtccactatgtagctctatgaacataataataatcctaattgagat 780  
DB 721 TTTCTGTAATATTCGTCACCTCATGACGTATGACACAAATTAATTAATCTTAATGGGAT 780  
QY 781 ttcaagtagaagaacgaggttaaaaagcttcgtgaagttactctaaagggaagaatcca 840  
DB 781 TTCAAGTAGAAGAACGGGGTTAAACGAGTCTGTCAGCTTAATCTTAAGGGAAAGTTCCA 840  
QY 841 gatgtgagccataactctgtccatctcaacagatgagctacaggaaggtgataacct 900  
DB 841 GATGTGAGCCATACCTGTGTCATTAACAGATGGCTATCAGGCAAGCAGGTGATACCT 900  
QY 901 cgtatgtgcttaattcatcaacagtaactatctatgaaggagatgacttgggaatacctat 960  
DB 901 CGTAAATGCTTAATTAATTCACACAGTACTACTATTCAGGAGATGATCTTGGGAATACCTAT 960  
QY 961 acacagaagaacaaacaccttaaaagcttggaacccaacttcaactaaatgacttctt 1020  
DB 961 ACACAGAAAGCAACACCTTAAGCTGTGGCACCAACTTCTACTCAAGTAATGTTCTT 1020  
QY 1021 cttatgacagtgacacggtgttcgtlaacaaaatcgtacctaagacggatcgggcat 1080  
DB 1021 CTTATGACAGTGACAAACGGGTTCTGTAAACAAAATCGTACCTATGACGGATGGGCCAT 1080  
QY 1081 ggtgtgtggaagcaacggttatacaaaaccttgaaaatggtatatacatgtatgagta 1140  
DB 1081 GGTGTGTGGGAAGCAACGGTTAATCAAAACCTTGAAAATTTGGATTAATACATGATGAGTA 1140  
QY 1141 acagaacgaagctctacccgaacggtgtgtatcctctatgacacacgcttgcacaaat 1200  
DB 1141 ACNAGGCAAGGCTCTACCCGAACGGCTGTGATCTTATGCAACTCGATTGCAACCAAT 1200  
QY 1201 ggaacgagagatgattgtgacacctgtgtlaaaacagatccctgtgcctggaacagtat 1260  
DB 1201 GGAACGAGAGGATGATGTGGACCTGGCTAAACAGATCTCTGCTGGCGAACAATGAT 1260  
QY 1261 aaacataataagccaagaatataagaagatgaggtacataagaatgagatgcgtgac 1320  
DB 1261 AAACATATTACGCAAAAGATATAGAAATGAGCTCATGTAATGATGTCCTGTGAC 1320  
QY 1321 ttttccatgacccctaatccggtatgtaaaaaataaagggaagatttgccttcaaga 1380  
DB 1321 TTTTCAATTGACCTTAATTCGGGTATGAAAATTAAGGGAAGATTGTGGCTTTACAGAA 1380  
QY 1381 aaaggaacaaagggccctgacaaacgtaaaagacggtgatactcttaaaacaacttgg 1440  
DB 1381 AAAGGAACAAAGGGCCCTGACAAAGTAAGACGGGATGATTCCTTAACAAACTTGGG 1440  
QY 1441 attaccatglttcagctttagctgttcttcgacatcaacagtgatgagtaaaactgtcca 1500  
DB 1441 ATTACTCATGTTCAAGCTTATGCTGTTCGCAATCAACAGTGTGATGAATGTATCCA 1500  
QY 1501 acccaagaataattgggttataagccctcgaaactgagatgctcgtlaaaggcgatgtgc 1560  
DB 1501 ACCCAAGATAATTGGGTTATAGCCTCGCAACTATGATGTTCTCTGAAGGGCATGTGCT 1560  
QY 1561 acaaatgcgaatgtgaaagctcgtataaaagagtttaaggaatggttcttcaactcat 1620  
DB 1561 ACAAATGCAAAATGTAATGCTCGTATAAAGAGTTTAAGGAATGGTTCTTCACTTCAT 1620  
QY 1621 cgtgaacacattgggttacaatgagtggttgcataatacatccttggccaagcaatc 1680  
DB 1621 CGTGAACACATTTGGGTTAAACATGAGTGTGTTAATATACCTTGGCCACGAAATC 1680  
QY 1681 tctgacttcgataaaaatttaccagatatattaccgtaacgtaacgtaagtaattata 1740  
DB 1681 TCTGACTTTCGATTAATTTGTACAGATATTTATACCGTACGATGATGACAGTAATTAAT 1740  
QY 1741 ccaacggtacaggtactggaataatgcaatgacangcgaaggccaatggttccaaaaatt 1800

Db 1741 CCAAGGATGAGTACGAAATCAATTCANCCNCAAGCCAAAGTTCACAAAATTT 1800  
 QY 1801 attatgattcccttaaglatltyggltcaatgatalcataltgaagctccgtlctgac 1860  
 Db 1801 ATTATGTATTCCTTAAGTANTGGCTCAATGAGATCATATTCAGCGCTTCCGTTTGAC 1860  
 QY 1861 ttaattgagctgcttggaaagacagatgltccaaagctgcctcgaagctcatgtact 1920  
 Db 1861 TTATGCGCGCTGTGGAAAAGACAGTGTCCAAAGCTGCGAGCTTCATGTATT 1920  
 QY 1921 aattccaaagattgcaacttcgtgagcctatgagcaggttgaaacctgtgacttgagat 1980  
 Db 1921 AATCCAGGAATGCACTTTAGGCTGACCCATGCAAGCGGTGAACCTCTGCACCTGCAGAT 1980  
 QY 1981 gatacagcttcgacaagaagagctcaaaaagcaltggaaglaagcgtgttaatgacaat 2040  
 Db 1981 GATCAGCTTCTGACAAAAGCAGCTCAAAAAGCATGGAGTAGCGGTTTAATGACAAT 2040  
 QY 2041 ttaagaaacgcttggacggaatgtcttgaatctcgcgtcaaggtlctgcaagat 2100  
 Db 2041 TTACGAACCGCTGGAGCGGAATGTCTTGATTCCTCCGCTCAAGGTTTGCGACAGGT 2100  
 QY 2101 gcaacaaagcttaacttgaatgaatgaatgaggttgaagaggtatgaatgaacttacc 2160  
 Db 2101 GCAACAGGCTTAAGTATGCAATTAAGAAATGCGCTTGAAGGAGTATTAATGACTTACC 2160  
 QY 2161 tcttcaaccaggtgagacaatlaactatgtcaagaatgaltgaactaacacctlctggac 2220  
 Db 2161 TCTTCACACAGGTGAGACAATTAAGTATGTCACAGCATGATTAACACCTTTGGGAC 2220  
 QY 2221 aaaaataagccctaaagcaatccttaagatctcgaagcagatcgatataaagtgaatc 2280  
 Db 2221 AAAATAGCCCTTAACCAATCCCAATGATTCGAAAGCGGATTAATGAATGATGAATC 2280  
 QY 2281 gcaacaaagcttgaatgaatcctcaagcgttcaactatgaagcaggtggaagaatg 2340  
 Db 2281 GCACAAAGCAGTCTTATGACTCTCAAGCGCTTCAATTCATGCAAGCGCGGAAGAAATG 2340  
 QY 2341 ctctgtaaaaaagcagcaacgacaatgaltatlaaagcagcgaatgcgtcaatgaat 2400  
 Db 2341 CTTCGTANAAAAAGCGGCAACGACAAATGATTAATGACAGCGCATGCAATGAGATT 2400  
 QY 2401 gattgagcaggaaagctcaatcaatcagatgttctcaactatgaagcgttaatccac 2460  
 Db 2401 GATTGGAGCAGGAAGCTCAATATCCAGATGTTTCAACTATTATAGCGGCTAATCCAC 2460  
 QY 2461 ctctgctctgaatcaacagccttcgcaatgaacgaacgaatgaatcaatgaacac 2520  
 Db 2461 CTTCGCTCTTATCAACCGCTTCCGATGACGACACTTAATCAATCAATAGCCACTC 2520  
 QY 2521 caatlcctlaaataagltccagaagacagatgagcclatgaatlaactgaatcgtlaataa 2580  
 Db 2521 CAATTCCTAAATACTCCAGAACACAGTGGCTATGCAATTAACCTGATCAATTAATAA 2580  
 QY 2581 gacaaatgaggaaatatactgttcttataaaccataaactgtgagcaaccataat 2640  
 Db 2581 GACAAATGGGAAATATCATTTGTTTATAACCAATAAAMCTGTAGCAACCATCAAT 2640  
 QY 2641 ttgacgaagcggaaatgagcaatcaatgtactaagcgaatgaagtagagaaatccac 2700  
 Db 2641 TTGCGCAGCGGAAATGCGCATATCAATGCAAGCGGTAAAGTAGAGAAATCCACCTT 2700  
 QY 2701 ggtcaagcagaaggaagltgltcaaglaacaglaatcatalgaatcctcatalcaag 2760  
 Db 2701 GGTCAAGCAGAGGAGAGTGTCAAGTACAGCTATATCTATGATGATTCCTTCATCAAGAG 2760  
 QY 2761 gtaagccagagaccaggttaaaag 2784  
 Db 2761 GTAAGCCAGACCAAGGTAAAG 2784

Search completed: April 22, 2001, 07:55:55  
 Job time: 5879 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2001, 06:11:02 ; Search time 174.28 Seconds  
(without alignments)  
9358.991 Million cell updates/sec

Title: US-09-262-126c-1

Perfect score: 2794  
Sequence: 1 gatgggaacacgacacgacat.....cggtaaaagtaataagaaa 2794

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_0401.\*  
1: /SID52/gcgdata/geneseq/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/NA2001.DAT.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2788	99.8	2794	20	Bacillus deramific
2	2786.4	99.7	4464	15	Bacillus deramific
3	2776.4	99.4	2784	15	Bacillus deramific
4	1823.6	65.3	3934	19	Trimmed enzyme enc
5	894.8	32.0	2736	21	Pullulanase (publ)
6	321.4	11.5	2371	13	Heat-resistant pul
7	276	9.9	2319	19	Thermotoga maritim
8	271.2	9.7	2319	18	Thermotoga maritim
9	269.6	9.6	2601	20	Fervidobacterium p
10	226	8.1	3683	16	Thermotoga maritim
11	171	6.1	9541	19	Streptococcus pneu

12	76.2	2.7	1993	18	T93875
13	76.2	2.7	2904	20	V65118
14	74.4	2.7	5058	20	X91106
15	73.6	2.6	784	20	X33179
16	72.8	2.6	3753	21	A05807
17	72.2	2.6	2646	16	T02405
18	72.2	2.6	2982	16	T02400
19	72.2	2.6	2988	16	T02399
20	68.2	2.4	3437	17	T31478
21	53.6	1.9	273254	21	C81914
22	53.6	1.9	1230025	20	X91990
23	52	1.9	2133	18	T96754
24	52	1.9	2706	20	X27061
25	49	1.8	4133	9	N81341
26	48	1.7	624	20	V65116
27	47.6	1.7	2244	17	T10429
28	47.6	1.7	2244	19	V23640
29	47.6	1.7	2334	17	T10428
30	47.6	1.7	2334	19	V23639
31	46	1.6	1860	21	A56836
32	45.6	1.6	2712	20	X76375
33	45.2	1.6	1663	20	X33162
34	45.2	1.6	2806	20	X27066
35	45	1.6	1038602	20	Z01425
36	43.4	1.6	1706	20	X34654
37	43.4	1.6	2235	10	N91089
38	43.4	1.6	3337	10	N91089
39	42.6	1.5	2997	21	Z24492
40	42.2	1.5	6142	17	T47878
41	42.2	1.5	6142	17	V62861
42	39.6	1.4	2949	21	D00024
43	39.4	1.4	2843	20	X27060
44	39.4	1.4	2843	20	X27065
45	39.4	1.4	16535	19	V52207

#### ALIGNMENTS

RESULT	1
XX	X87941
XX	22-NOV-1999 (first entry)
AC	X87941; standard; DNA: 2794 BP.
XX	
DE	Bacillus deramificans mature pullulanase coding sequence.
XX	
KW	pullulanase; enzyme engineering; starch; saccharification;
KW	straight chain amylase; alpha-dextrin 6-glucano-hydrolase;
KW	alpha-1,6-glucosidase; ss.
XX	
OS	Bacillus deramificans.
XX	
PH	Key
FT	1..2787
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/EC_number= "3.2.1.41"
FT	/Product= "pullulanase"
XX	
PN	W09945124-A2.
XX	
PD	10-SEP-1999.
XX	
PF	03-MAR-1999; 99WO-US04627.
XX	
PR	04-MAR-1998; 98US-0034630.
XX	
PA	(GENV ) GENENCOR INT INC.
XX	
PI	Miller BS, Shetty JK;
XX	

2. mays debranchin  
Corn pullulanase c  
Group B Streptococ  
pullulanase type par  
Group B Streptococ  
Rice mature pullul  
pullulanase expres  
Rice mature pullul  
Spinach debranchin  
Chlamydia pneumoni  
Nucleotide sequenc  
S. tuberosum debra  
S. tuberosum isom  
Recombinant plasm  
Corn pullulanase c  
Mature isomylase  
Flavobacterium iso  
isoamylase gene.  
Full length Flavob  
mycoplasma hyopenu  
zea mays SU1 starch  
Potato isomylase  
Original S. tubero  
Complete genome se  
Partial cDNA for h  
DNA encoding polyp  
Fragment of Pseudo  
wheat isomylase c  
Bacillus alkaline  
Full length Pullul  
Degenerate DNA enc  
S. tuberosum isom  
Original S. tubero  
Streptococcus pneu

DR WP1: 1999-540851/45.  
DR P-PSDB: Y31746.

XX New modified pullulanase for saccharification of aqueous liquefied  
PT starch  
XX

PS Claim 18: Fig 1A-E: 49pp: English.

XX The present sequence represents *Bacillus deramificans* strain  
CC 789.117D DNA coding for mature pullulanase (EC-3.2.1.41). This  
CC enzyme has a pI of 4.1-4.5, has a half-life of about 55 hr at 60  
CC deg C and is capable of catalysing the hydrolysis of  
CC alpha-1,6-glucosidic bonds in amylopectin and pullulan. The  
CC invention relates to the discovery that *B. deramificans* pullulanase  
CC recombinantly produced in a *Bacillus* host is modified yet retains  
CC the ability to catalyse the hydrolysis of an alpha-1,6-glucosidic  
CC bond. The modification of the recombinant pullulanase appears to  
CC be a result of misprocessing of the signal sequence by a signal  
CC peptidase as well as susceptibility to extracellular proteolytic  
CC processing. The invention provides modified pullulanases,  
CC nucleic acids encoding them, vectors, host cells and methods for  
CC recombinant production. The modified pullulanases are useful in  
CC the starch industry for the saccharification of aqueous liquefied  
CC starch, forming syrups of high glucose and maltose content. Other  
CC uses (not claimed) are: as anti-staling additives in bread; in  
CC brewing low calorie beers; production of low calorie foods (where  
CC amylose is used as substitute for fat); to clarify fruit juices;  
CC In the preparation of oligosaccharides from amylopectin or of  
CC tetraholosides from maltose; to condense mono- or oligosaccharides,  
CC or in the pharmaceutical industry (no details). When used for  
CC starch saccharification, the modified pullulanases reduce the  
CC amounts of by-products, especially they provide higher glucose  
CC yields in short reaction times, without losses associated with  
CC reversal reaction products. When used with glucoamylase, they  
CC make possible a higher level of dissolved solids in the material  
CC being treated, resulting in increased productivity and a more  
CC energy-efficient process.

XX Sequence 2794 BP: 871 A; 559 C; 635 G; 723 T; 6 other;

Query Match 99.8% Score 2788; DB 20; Length 2794;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatvggaacacgacacgacatcatgtccactattctgcctcgtcgtgattcaacct 60  
DB 1 gatvggaacacgacacgacatcatgtccactattctgcctcgtcgtgattcaacct 60  
OY 61 tggagctatagatgtggccaaagaggggtgggctgaatcgaattcaatcaacg 120  
DB 61 tggagctatagatgtggccaaagaggggtgggctgaatcgaattcaatcaacg 120  
OY 121 gclgaaccltttgagcgtgtgcaagtcgataatccagaagaaacccaagtcagtaga 180  
DB 121 gclgaaccltttgagcgtgtgcaagtcgataatccagaagaaacccaagtcagtaga 180  
OY 121 gclgaaccltttgagcgtgtgcaagtcgataatccagaagaaacccaagtcagtaga 180  
DB 121 gclgaaccltttgagcgtgtgcaagtcgataatccagaagaaacccaagtcagtaga 180  
OY 181 attatcgtcgcactcaagatltggacaagaatgtgacgctgacgctacatagattta 240  
DB 181 attatcgtcgcactcaagatltggacaagaatgtgacgctgacgctacatagattta 240  
OY 241 agcaaaagaaatggagtggtgtgtgtagaagaaacagcaaattttttaatagaanaa 300  
DB 241 agcaaaagaaatggagtggtgtgtgtagaagaaacagcaaattttttaatagaanaa 300  
OY 301 gatgcctgaagatgcagctaaacccgcgttaagcaacgctattatagatgctcaaacag 360  
DB 301 gatgcctgaagatgcagctaaacccgcgttaagcaacgctattatagatgctcaaacag 360  
OY 361 gtgcgtgttaaaacttagccagcgctaaactcttgggggaagnnaaagcgcttaacgct 420  
DB 361 gtgcgtgttaaaacttagccagcgctaaactcttgggggaagnnaaagcgcttaacgct 420

OY 421 catgacgacacgacgaataatagatattccagtgacatctgtgaagatgtgaactcgtc 480  
DB 421 catgacgacacgacgaataatagatattccagtgacatctgtgaagatgtgaactcgtc 480  
OY 481 caagatgtaacccgtgttctggcaggtacccccaacataatttttggaggtccgattgg 540  
DB 481 caagatgtaacccgtgttctggcaggtacccccaacataatttttggaggtccgattgg 540  
OY 541 gcaactgtataatcaacgtactattataaaaaaggtgatacaaatctctataatctca 600  
DB 541 gcaactgtataatcaacgtactattataaaaaaggtgatacaaatctctataatctca 600  
OY 601 gggagatcttcctgaagaaactaccataataaagtgtccttaaatgtatgaatgaat 660  
DB 601 gggagatcttcctgaagaaactaccataataaagtgtccttaaatgtatgaatgaat 660  
OY 661 ccagatccacactgtgacacatattataacgcttcctcggcggtgacacgctact 720  
DB 661 ccagatccacactgtgacacatattataacgcttcctcggcggtgacacgctact 720  
OY 721 tttcgtatattccgttcactcatgtcgtatgtacacaaattataatcttaatgcygat 780  
DB 721 tttcgtatattccgttcactcatgtcgtatgtacacaaattataatcttaatgcygat 780  
OY 781 tttaacgtataaagcggtgttaaaacggtcgtgtgacgtttctctaggggaaatcca 840  
DB 781 tttaacgtataaagcggtgttaaaacggtcgtgtgacgtttctctaggggaaatcca 840  
OY 841 gatgtgacacatactcgttcacatccaacagatgtgcatacgaacgaagcggtgatact 900  
DB 841 gatgtgacacatactcgttcacatccaacagatgtgcatacgaacgaagcggtgatact 900  
OY 901 cgtaatgtccttaattcaatcaacgtactactatcagagagatgattcttgggaatacct 960  
DB 901 cgtaatgtccttaattcaatcaacgtactactatcagagagatgattcttgggaatacct 960  
OY 961 acaagaagaacaaac 1020  
DB 961 acaagaagaacaaac 1020  
OY 1021 cttatgacagctgcacacgggtctcgttaacaaacacacacacacacacacacacacac 1080  
DB 1021 cttatgacagctgcacacgggtctcgttaacaaacacacacacacacacacacacacac 1080  
OY 1081 ggtgtgtgggaagcaacggttaatacaaaccttgaaatttgtaatacgtatgagta 1140  
DB 1081 ggtgtgtgggaagcaacggttaatacaaaccttgaaatttgtaatacgtatgagta 1140  
OY 1141 acaaggaagacgctcctacacgaacggctgtgtgacacacacacacacacacacacacac 1200  
DB 1141 acaaggaagacgctcctacacgaacggctgtgtgacacacacacacacacacacacacac 1200  
OY 1201 ggaacgaagacatgattgtgacacgtgtgacacacacacacacacacacacacacac 1260  
DB 1201 ggaacgaagacatgattgtgacacgtgtgacacacacacacacacacacacacacac 1260  
OY 1261 aaacataatcgcacaaagaatatagaagatgaggtcctcctatgaaatgtgtcgtgac 1320  
DB 1261 aaacataatcgcacaaagaatatagaagatgaggtcctcctatgaaatgtgtcgtgac 1320  
OY 1321 ttttcctatgaac 1380  
DB 1321 ttttcctatgaac 1380  
OY 1381 aaaggaacaaagggccctcgaacacacacacacacacacacacacacacacacacacac 1440  
DB 1381 aaaggaacaaagggccctcgaacacacacacacacacacacacacacacacacacacac 1440  
OY 1441 attactcatgtcaagctatgctgttctcgcacatacaacggtcgtcgtgaacatgataca 1500  
DB 1441 attactcatgtcaagctatgctgttctcgcacatacaacggtcgtcgtgaacatgataca 1500  
OY 1501 acccaagataattggggtttagacccctgcacacatgatttctcgtgaagggcgtatgct 1560



DR P-PSDB; R56990.

XX. New pullulanase from *Bacillus deramificans* - for starch  
PT saccharification etc., with good stability over wide temp. and pH  
PT ranges, also related DNA vectors, transformed cells etc.  
XX

PS Claim 10; Fig 5; 61pp; French.

XX This sequence codes for the pullulanase enzyme precursor isolated  
CC from *Bacillus deramificans* T 89.117D. The pullulanase hydrolyses the  
CC alpha-1,6-glycosidic bonds of amylopectin and of pullulan. The  
CC enzyme can be used (with glucoamylase) for saccharification of  
CC starch. Typical applications include use as anti-staling additives  
CC in bread making and brewing, in preparation of low calorie foods  
CC to clarify fruit juices.  
XX

SQ Sequence 4464 BP; 1355 A; 855 C; 1039 G; 1187 T; 28 other;

Query Match 99.7%; Score 2786.4; DB 15; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggagacacgacacacatcattgtccactattttgacctgtgtgtgattacaacct 60  
|||||  
Db 1250 gatggagacacgacacacatcattgtccactattttgacctgtgtgtgattacaacct 1309  
QY 61 tggagtcattgagatgtggcacaagaacgaggtggggctggaatacgaattccaatcaacg 120  
|||||  
Db 1310 tggagtcattgagatgtggcacaagaacgaggtggggctggaatacgaattccaatcaacg 1369  
QY 121 ggcgaccttttggagctgtgttcgaagctgcatattccagaagaaacccaagtcagtgaga 180  
|||||  
Db 1370 ggcgaccttttggagctgtgttcgaagctgcatattccagaagaaacccaagtcagtgaga 1429  
QY 181 attatcctgcgcctcacaagatttgacacaagaatgtgagcgtcagcgtccatagattta 240  
|||||  
Db 1430 attatcctgcgcctcacaagatttgacacaagaatgtgagcgtcagcgtccatagattta 1489  
QY 241 agcacaagaaatgagatgtgtgtgtgtgaagaagaacacgaatttttaataagaaaaa 300  
|||||  
Db 1490 agcacaagaaatgagatgtgtgtgtgtgaagaagaacacgaatttttaataagaaaaa 1549  
QY 301 gatgctgtagagatgtcagcttaaacccgcgttaagcaacgcttattagctgtccaaacag 360  
|||||  
Db 1550 gatgctgtagagatgtcagcttaaacccgcgttaagcaacgcttattagctgtccaaacag 1609  
QY 361 gtcctggttaaaacttaacgacgcttaactcttgaggaaagmnaagcggcttaacggtc 420  
|||||  
Db 1610 gtcctggttaaaacttaacgacgcttaactcttgaggaaagmnaagcggcttaacggtc 1669  
QY 421 catgacgacacagcaaaataagatuttcacgtgacatctgttgaaagatgcaagtcttgt 480  
|||||  
Db 1670 catgacgacacacaaataagatuttcacgtgacatctgttgaaagatgcaagtcttgt 1729  
QY 481 caagagatcaacccgtgttttggcaggtacacttccaaatatttttggaggtccogatttg 540  
|||||  
Db 1730 caagagatcaacccgtgttttggcaggtacacttccaaatatttttggaggtccogatttg 1789  
QY 541 gcaacctgataatcacagctacttatttaaaaaaggctgaatacactctctcaatctca 600  
|||||  
Db 1790 gcaacctgataatcacagctacttatttaaaaaaggctgaatacactctctcaatctca 1849  
QY 601 gggagatcttcctgaaagaaatacaataaagtgtgcttaataatgtagagctggaaat 660  
|||||  
Db 1850 gggagatcttcctgaaagaaatacaataaagtgtgcttaataatgtagagctggaaat 1909  
QY 661 ccgaagtaaccatctggacaacatttaatttaacagctccctcggcggtgacacgcttact 720  
|||||  
Db 1910 ccgaagtaaccatctggacaacatttaatttaacagctccctcggcggtgacacgcttact 1969  
QY 721 ttctctgataatccgtccacatctgcagctcagacaaatlaataatccaaatgacgat 780  
|||||

Db 1970 ttctctgataatccgtccacatctgcagctcagacaaatlaataatccaaatgacgat 2029  
QY 781 ttacaagtaagaagcgggtgttaaaacgagatctctgacggttactctagggaagaccca 840  
|||||  
Db 2030 ttacaagtaagaagcgggtgttaaaacgagatctctgacggttactctagggaagaccca 2089  
QY 841 gatgtgacatctctgtccatctcaaacagatggtcatalcaggaacaggtgatcact 900  
|||||  
Db 2090 gatgtgacatctctgtccatctcaaacagatggtcatalcaggaacaggtgatcact 2149  
QY 901 cgttaattgttttaattatataacgactactattctcaggaatgatctcttggaaactat 960  
|||||  
Db 2150 cgttaattgttttaattatataacgactactattctcaggaatgatctcttggaaactat 2209  
QY 961 acacagaagaacaaacacttataagctgtggcacaacactctacacgaatgaattgtctc 1020  
|||||  
Db 2210 acacagaagaacaaacacttataagctgtggcacaacactctacacgaatgaattgtctc 2269  
QY 1021 cttaatacagctgcacaggtctctgttaacaaaatcgttacctatgacgacatcggtccat 1080  
|||||  
Db 2270 cttaatacagctgcacaggtctctgttaacaaaatcgttacctatgacgacatcggtccat 2329  
QY 1081 ggtgtgtggaagaacggttaatacacaacacttggaataatgtgatatcagatgaggtta 1140  
|||||  
Db 2330 ggtgtgtggaagaacggttaatacacaacacttggaataatgtgatatcagatgaggtta 2389  
QY 1141 acagggcacaagctctacacgacggtctgtatctcttaatgcaactcgtgattgacacaaat 1200  
|||||  
Db 2390 acagggcacaagctctacacgacggtctgtatctcttaatgcaactcgtgattgacacaaat 2449  
QY 1201 ggaacgagaagcattgattgtggacactgtgttaaaaacagatccctgtgctgtgaacagtat 1260  
|||||  
Db 2450 ggaacgagaagcattgattgtggacactgtgttaaaaacagatccctgtgctgtgaacagtat 2509  
QY 1261 aaaaattatcgcccaaaagaatatagaagatgaggtctctcttaaaatggatgtgtccgtgac 1320  
|||||  
Db 2510 aaaaattatcgcccaaaagaatatagaagatgaggtctctcttaaaatggatgtgtccgtgac 2369  
QY 1321 ttctccattgacccctaaattcgggttatgaaataataaagggaagatcttgcgtcttaacaga 1380  
|||||  
Db 2570 ttctccattgacccctaaattcgggttatgaaataataaagggaagatcttgcgtcttaacaga 2629  
QY 1381 aaaaagacaaagggccctgtgacacgttaaaagcgggtgataatcccttaaaaacacttggtg 1440  
|||||  
Db 2630 aaaaagacaaagggccctgtgacacgttaaaagcgggtgataatcccttaaaaacacttggtg 2689  
QY 1441 attactcatgtcagctctgacgtgttttcgcatacctcaacagcttctgagtgaaactgtatcca 1500  
|||||  
Db 2690 attactcatgtcagctctgacgtgttttcgcatacctcaacagcttctgagtgaaactgtatcca 2749  
QY 1501 acccaagataatctgggttataagacccctgcgaactatgagtctcctgaaaggcagatagct 1560  
|||||  
Db 2750 acccaagataatctgggttataagacccctgcgaactatgagtctcctgaaaggcagatagct 2809  
QY 1561 accaaatgtcgaatgttaattgtctgtatataaagaatttaagaagaatgttcttcaatccat 1620  
|||||  
Db 2810 accaaatgtcgaatgttaattgtctgtatataaagaatttaagaagaatgttcttcaatccat 2869  
QY 1621 cgttgaacacacttgggttataacatgagtgtctcatataatcccttggcagcgaacatc 1680  
|||||  
Db 2870 cgttgaacacacttgggttataacatgagtgtctcatataatcccttggcagcgaacatc 2929  
QY 1681 tctgactctgataaaatctgacagaatattatccagtgacagatgacagatgataata 1740  
|||||  
Db 2930 tctgactctgataaaatctgacagaatattatccagtgacagatgacagatgataata 2989  
QY 1741 ccgaagatacaggttactggaatgtgaaattgacangcngaaagccaaatggttcaaaaattc 1800  
|||||  
Db 2990 ccgaagatacaggttactggaatgtgaaattgacangcngaaagccaaatggttcaaaaattc 3049  
QY 1801 attatgactcccttaagatactgtgtcaatgagatcaatacttgaagcttcggtttgac 1860  
|||||  
Db 3050 attatgactcccttaagatactgtgtcaatgagatcaatacttgaagcttcggtttgac 3109



```

OY 1861 ttaatgagcgtcttggaagaaagacagatgiccagaagctgcctcggagcttcattatc 1920
    |||||||
DB 3110 ttaatgagcgtcttggaagaaagacagatgiccagaagctgcctcggagcttcattatc 3169
    |||||||
OY 1921 aatccagaatctgacatttaagtgtagcagatgagcgggttggaacctcgcacgtccaat 1980
    |||||||
DB 3170 aatccagaatctgacatttaagtgtagcagatgagcgggttggaacctcgcacgtccaat 3229
    |||||||
OY 1981 gatcagcttcgacaaaagagagcctcaaaaagccttgagagtagcggtgtttaataacaaat 2040
    |||||||
DB 3230 gatcagcttcgacaaaagagagcctcaaaaagccttgagagtagcggtgtttaataacaaat 3289
    |||||||
OY 2041 ttaagaaacgcttgagacggaatgcttcttgatcttcctgcgtcgaaggttttcgacaggt 2100
    |||||||
DB 3290 ttaagaaacgcttgagacggaatgcttcttgatcttcctgcgtcgaaggttttcgacaggt 3349
    |||||||
OY 2101 gcaacagcgttaactgtagcaataaagatgagcgttgagggagtagtaataatgactttacc 2160
    |||||||
DB 3350 gcaacagcgttaactgtagcaataaagatgagcgttgagggagtagtaataatgactttacc 3409
    |||||||
OY 2161 tcttcacacaggtgagacaaatgaatgaatgacagatgacatgatacaccctttgggac 2220
    |||||||
DB 3410 tcttcacacaggtgagacaaatgaatgaatgacagatgacatgatacaccctttgggac 3469
    |||||||
OY 2221 aaaaatagccttaagcaatcctaataatgattcgaagcggaatcgatataaattgatgaactc 2280
    |||||||
DB 3470 aaaaatagccttaagcaatcctaataataatgattcgaagcggaatcgatataaattgatgaactc 3529
    |||||||
OY 2281 gcaacagcagttgttataatgacttcacaaagcgttccatcagacagcgagggaagaaatg 2340
    |||||||
DB 3530 gcaacagcagttgttataatgacttcacaaagcgttccatcagacagcgagggaagaaatg 3589
    |||||||
OY 2341 ctccgtanaaagacgagcaacagacatagttataatgacggcgatcgagcgttcgaatgaatt 2400
    |||||||
DB 3580 ctccgtanaaagacgagcaacagacatagttataatgacggcgatcgagcgttcgaatgaatt 3649
    |||||||
OY 2401 gattggagcaggaagatcaatataatccagaatgtttcaactaattatagcgggctaatac 2460
    |||||||
DB 3650 gattggagcaggaagatcaatataatccagaatgtttcaactaattatagcgggctaatac 3709
    |||||||
OY 2461 ctctgcttctgatacccccgccttcgcgacatgacagatgaatgaataatgaacacatc 2520
    |||||||
DB 3710 ctctgcttctgatacccccgccttcgcgacatgacagatgaatgaataatgaacacatc 3769
    |||||||
OY 2521 caatccctaataatgacagagaacacagatggcctatgaattaactgacatgatttaataa 2580
    |||||||
DB 3770 caatccctaataatgacagagaacacagatggcctatgaattaactgacatgatttaataa 3829
    |||||||
OY 2581 gacaaatggggaatatatcatgtgttataaaccacaataaactgtgacaaacatcaat 2640
    |||||||
DB 3830 gacaaatggggaatatatcatgtgttataaaccacaataaactgtgacaaacatcaat 3889
    |||||||
OY 2641 ttgcccgaagcgggaaatgtagcaatcaatgctagagcggttaaggttgaggaatccaccctt 2700
    |||||||
DB 3890 ttgcccgaagcgggaaatgtagcaatcaatgctagagcggttaaggttgaggaatccaccctt 3949
    |||||||
OY 2701 ggtcaagcagcggggaaggtgtccaagtaccaggtatataatgatattatcattatcaagaag 2760
    |||||||
DB 3950 ggtcaagcagcggggaaggtgtccaagtaccaggtatataatgatattatcattatcaagaag 4009
    |||||||
OY 2761 gtaagccagacacacggttaaaaaatagaaaa 2794
    |||||||
DB 4010 gtaagccagacacacggttaaaaaatagaaaa 4043
    |||||||

```

```

RESULT 3
ID Q68698 standard: DNA: 2784 BP.
XX Q68698:
AC
XX
DT 20-FEB-1995 (first entry)

```

```

XX XX Bacillus deramificans mature pullulanase genomic DNA.
DE DE
XX XX pullulanase; recombinant production; pullulan; amylopectin;
KW KW hydrolysis; alpha-1,6-glycosidic bond; starch saccharification; ss.
XX XX
OS Bacillus deramificans T 89.117D (LMG P-13056).
XX OS
XX XX
FH Key
FH mat_peptide
FH 1..2784
FH /tag- a
FH /product- mature-pullulanase
FH 402..404
FH /tag- b
FH /note- "GCN NNA encodes Gly-Xaa"
FH 1773
FH /tag- c
FH /note- "CAN encodes His or Gln"
FH 1776
FH /tag- d
FH /note- "GCN encodes Ala"
FH 2348
FH /tag- e
FH /note- "ANA encodes Xaa"
FH
XX EP605040-A.
XX
XX PD 06-JUL-1994.
XX
XX PF 20-DEC-1993; 93EP-0203593.
XX
XX PR 28-DEC-1992; 92BE-0001156.
XX PR 15-JUL-1993; 93BE-0000744.
XX PR 19-NOV-1993; 93BE-0001278.
XX
XX PA (SOLV ) SOLVAY SA.
XX
XX PI Amory A, DeWeer P;
XX
XX DR MPI: 1994-210291/26.
XX
XX DR P-PSDB: R56989.
XX
XX CC This sequence codes for the mature pullulanase enzyme isolated from
XX CC Bacillus deramificans T 89.117D. The pullulanase hydrolyses the
XX CC alpha-1,6-glycosidic bonds of amylopectin and of pullulan. The
XX CC enzyme can be used (with glucoamylase) for saccharification of
XX CC starch. Typical applications include use as anti-staling additives
XX CC in bread making and brewing, in preparation of low calorie foods
XX CC and to clarify fruit juices.
XX
SQ Sequence 2784 BP; 864 A; 558 C; 635 G; 721 T; 6 other:

```

```

Query Match 99.4%; Score 2776.4; DB 15; Length 2784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 gatggaaacagacacagatcatgtccacatattttgcccgtcgtgattatacaact 60
    |||||||
DB 1 gatggaaacagacacagatcatgtccacatattttgcccgtcgtgattatacaact 60
    |||||||
OY 61 tggagctctatgtagtcttgccaaagacgaggttgaggctgtaatacattcaacaaccg 120
    |||||||
DB 61 tggagctctatgtagtcttgccaaagacgaggttgaggctgtaatacattcaacaaccg 120
    |||||||
OY 121 gctgactcttttgaggcgtgttcgaagtgtgtatattccaggaaaccaagtcaagttagga 180
    |||||||

```

D	b	121	gctgacccctcttgagagctgtgtctgaagtgctgtatattccaggaagaaacccaagctcagtgtaga	180
Q	y	181	atatacgtctcgccctccaaagatctggaaccaaagatgtagagcgctcgacccgtatacatgattta	240
D	b	181	attatcgtctcgccctccaaagatctggaaccaaagatgtagagcgctcgacccgtatacatgattta	240
Q	y	241	agcaaaagaaaatgtaggtgtgtgctctgtatagaagagaaacagccaaattttttataatgaaaaa	300
D	b	241	agcaaaagaaaatgtaggtgtgtgctctgtatagaagagaaacagccaaattttttataatgaaaaa	300
Q	y	301	gatctctaggaatccagctcaaaccccgctgtgtaagacaacgctattcttagatgcttcaaaacag	360
D	b	301	gatctctaggaatccagctcaaaccccgctgtgtaagacaacgctattcttagatgcttcaaaacag	360
Q	y	361	gtgcctggtctaaacttaagccagccgtttaacctctgtggaaagmnaaacgsgctttaaagtt	420
D	b	361	gtgcctggtctaaacttaagccagccgtttaacctctgtggaaagmnaaacgsgctttaaagtt	420
Q	y	421	catgtagcaacacagcaataaagatattccagatgacacatctgtgaaagatgcaagctctggt	480
D	b	421	catgtagcaacacagcaataaagatattccagatgacacatctgtgaaagatgcaagctctggt	480
Q	y	481	caagatctaaacccgtctgtgtggcaggctaacctcccaacatcttctggaggtctccagatgg	540
D	b	481	caagatctaaacccgtctgtgtggcaggctaacctcccaacatcttctggaggtctccagatgg	540
Q	y	541	gcaccctgaataatccacagtaacttatttaaaaaggtgagctaaacatctctatcaattctca	600
D	b	541	gcaccctgaataatccacagtaacttatttaaaaaggtgagctaaacatctctatcaattctca	600
Q	y	601	ggagatctctctgtgaaggaatactccaatataaagtgtctttaaatgatacagcttgaaataat	660
D	b	601	ggagatctctctgtgaaggaatactccaatataaagtgtctttaaatgatacagcttgaaataat	660
Q	y	661	ccgsggtctaaccaatctgtgacaacatcttaatacagctccctcgccgsggtgtgacaacgtcaat	720
D	b	661	ccgsggtctaaccaatctgtgacaacatcttaatacagctccctcgccgsggtgtgacaacgtcaat	720
Q	y	721	ttctctgataatctcgttccacttcatgtcagctctatgaaacaaataataatcaatcttaatg	780
D	b	721	ttctctgataatctcgttccacttcatgtcagctctatgaaacaaataataatcaatcttaatg	780
Q	y	781	tttaacagatgagaagcggtgttaaaaacagatctccgtgtgacggttaactctctggtggaaatcca	840
D	b	781	tttaacagatgagaagcggtgttaaaaacagatctccgtgtgacggttaactctctggtggaaatcca	840
Q	y	841	gattgtgagccaaatctgtcccatctcaaacagatagtgctacaggaabaagcaggtgtgatact	900
D	b	841	gattgtgagccaaatctgtcccatctcaaacagatagtgctacaggaabaagcaggtgtgatact	900
Q	y	901	cgtaaatgtgcttaaatctcatcacagtaactatctcaagagatgatactcttgggaataactcat	960
D	b	901	cgtaaatgtgcttaaatctcatcacagtaactatctcaagagatgatactcttgggaataactcat	960
Q	y	961	acacagagaagacaaacacctttaaagctctgtggcaccaaacttctaactcaagtaaatgctctt	1020
D	b	961	acacagagaagacaaacacctttaaagctctgtggcaccaaacttctaactcaagtaaatgctctt	1020
Q	y	1021	ctttatgacacgttgacaaggggtctctgttaacaaaatctgtataccatgtagcggtcaggtccat	1080
D	b	1021	ctttatgacacgttgacaaggggtctctgtgttaacaaaatctgtataccatgtagcggtcaggtccat	1080
Q	y	1081	gggtgtgtggaagaaagaaacggtttaaataaaacaccttgaaatctgttatacatgataaggtta	1140
D	b	1081	gggtgtgtggaagaaagaaacggtttaaataaaacaccttgaaatctgttatacatgataaggtta	1140
Q	y	1141	acagagccaaagctcttaacccgaacgctgtgtatcccttaatgcaactgtgacatctgacccaat	1200
D	b	1141	acagagccaaagctcttaacccgaacgctgtgtatcccttaatgcaactgtgacatctgacccaat	1200
Q	y	1201	ggaaacggagagcatgatatgtgtgacacccgtgctataaaacagatccctgctgtgacaaagtat	1260
D	b	1201	ggaaacggagagcatgatatgtgtgacacccgtgctataaaacagatccctgctgtgacaaagtat	1260

QY	1261	aaacatacttcgcgcacaaagaaataataagaaagcttgagcgcaacttatgaaatgtagatgctccgagac	1320
Db	1261	aaacatacttcgcgcacaaagaaataataagaaagcttgagcgctacattatgaaataagatgtagctccgagac	1320
QY	1321	ttctccatactgacccctaaatctcggtatgaaataataataagaaagaaatattctgaccttcacgaag	1380
Db	1321	ttctccatactgacccctaaatctcggtatgaaataataataagaaagaaatattctgaccttcacgaag	1380
QY	1381	aaaggaacaaagggcccttgacaaagcttaaaagacgagggaatagatctccctaaataaacaacttgag	1440
Db	1381	aaaggaacaaagggcccttgacaaagacggaagagcgaggatgtagattctccctaaataaacaacttgag	1440
QY	1441	attctcatactgctcaagccttaagcctgtctttccgcatactcaacaagagctgcagaaacgagatcca	1500
Db	1441	attctcatactgctcaagccttaagcctgtctttccgcatactcaacaagagctgcagaaacgagatcca	1500
QY	1501	accacaaagataacttgagggtttaagacccctcgcaacatactgattctccctgaagaggcgatagct	1560
Db	1501	accacaaagataacttgagggtttaagacccctcgcaacatactgattctccctgaagaggcgatagct	1560
QY	1561	acaaatgcgaaatgtagaattgctgcgtataaaaagagtttaagggaaatgtagttctcttcacatccat	1620
Db	1561	acaaatgcgaaatgtagaattgctgcgtataaaaagagtttaagggaaatgtagttctcttcacatccat	1620
QY	1621	cgttgaaacaacacttgagggtttcaaacatcgagatgctgtctataataacataccactctgcacgaacac	1680
Db	1621	cgttgaaacaacacttgagggtttcaaacatcgagatgctgtctataataacataccactctgcacgaacac	1680
QY	1681	tcgtgactctgataaaatattgtaacccaagaataattattacacgttaagaatgtagccgaagtataata	1740
Db	1681	tcgtgactctgataaaatattgtaacccaagaataattattacacgttaagaatgtagccgaagtataata	1740
QY	1741	ccaaacgagatccagtgatccgagaaatgaaatattgcgaagcgaagaaagccgaatggttcacaaaatct	1800
Db	1741	ccaaacgagatccagtgatccgagaaatgaaatattgcgaagcgaagaaagccgaatggttcacaaaatct	1800
QY	1801	attatctgactcccttaagctatctgaggttcaaatgtagatcabaatattgacgctctccgattctgag	1860
Db	1801	attatctgactcccttaagctatctgaggttcaaatgtagatcabaatattgacgctctccgattctgag	1860
QY	1861	ttaatgtagcgcgtcttgagaaagaaagacagatagctccaaagacgcgccttcgagacttcacatgacat	1920
Db	1861	ttaatgtagcgcgtcttgagaaagaaagacagatagctccaaagacgcgccttcgagacttcacatgacat	1920
QY	1921	aatccaaagaaatctgacactcttaacggctgaagccatctgagacggggttgagaaacctccgacactcgcagat	1980
Db	1921	aatccaaagaaatctgacactcttaacggctgaagccatctgagacggggttgagaaacctccgacactcgcagat	1980
QY	1981	gatacagctcttcgcaacaaagaaagcttcacaaagagcgaatgagggatgtagcggtttataagacaaat	2040
Db	1981	gatacagctcttcgcaacaaagaaagcttcacaaagagcgaatgagggatgtagcggtttataagacaaat	2040
QY	2041	ttacgaaagacgcgtttcgagccgagcaatctgactctgattctcccgacccaagagcttttcgcaacggt	2100
Db	2041	ttacgaaagacgcgtttcgagccgagcaatctgactctgattctcccgacccaagagcttttcgcaacggt	2100
QY	2101	gcaacacagccttaacctgatactgaatlaagaaatgtagcgtttgaagggagatattaaatgacttaac	2160
Db	2101	gcaacacagccttaacctgatactgaatlaagaaatgtagcgtttgaagggagatattaaatgacttaac	2160
QY	2161	tccttcacacaggtgagaaataataactatgtaacaaagatgataaataaataaaccctttggagac	2220
Db	2161	tccttcacacaggtgagaaataataactatgtaacaaagatgataaataaataaaccctttggagac	2220
QY	2221	aaaaatagcccttaagcaaatcccttaatactgattccgaagagcgatattcaaaaatgtagatgaacac	2280
Db	2221	aaaaatagcccttaagcaaatcccttaatactgattccgaagagcgatattcaaaaatgtagatgaacac	2280
QY	2281	gcacaaagcagttgtatlgaccttcacaaagcgttccatctcatgcaagagcgagggaagaaatg	2340
Db	2281	gcacaaagcagttgtatlgaccttcacaaagcgttccatctcatgcaagagcgagggaagaaatg	2340

QY 2341 ctctgtanaaaagcgcgaacagacatagtatataatcagcgatcggtcaatgaattt 2400  
DB 2341 ctctgtanaaaagcgcgaacagacatagtatataatcagcgatcggtcaatgaattt 2400  
QY 2401 gattggagcaggaagcccaatccacagatgtttcaactatattagcggttaaccac 2460  
DB 2401 gattggagcaggaagcccaatccacagatgtttcaactatattagcggttaaccac 2460  
QY 2461 ctctgtctgttccacccagccttccgcatacgacagcaatgaatcaatcaaccctc 2520  
DB 2461 ctctgtctgttccacccagccttccgcatacgacagcaatgaatcaatcaaccctc 2520  
QY 2521 caattcctaataatagtcacagagaacacagtcgacctatgaatactgacatgttaataa 2580  
DB 2521 caattcctaataatagtcacagagaacacagtcgacctatgaatactgacatgttaataa 2580  
QY 2581 gacaaatgggaaatataatgttgtttataaccacaaatgaactgtgcaaccatcaat 2640  
DB 2581 gacaaatgggaaatataatgttgtttataaccacaaatgaactgtgcaaccatcaat 2640  
QY 2641 ttgcgcagcgggaaatgggcaatcagtcacagcgttaagtgtaggaatccaccctt 2700  
DB 2641 ttgcgcagcgggaaatgggcaatcagtcacagcgttaagtgtaggaatccaccctt 2700  
QY 2701 ggtcaagcagcgggaaatgggcaatcagtcacagcgttaatctatctatctatcaagaag 2760  
DB 2701 ggtcaagcagcgggaaatgggcaatcagtcacagcgttaatctatctatctatcaagaag 2760  
QY 2761 gtaagccagacacacggttaaaaag 2784  
DB 2761 gtaagccagacacacggttaaaaag 2784

## RESULT 4

T97233  
ID T97233 standard; DNA: 3934 BP.

XX AC T97233;

XX DT 11-MAR-1998 (first entry)

XX DE Trimmed enzyme encoding DNA.

XX KW Bacillus sp. APC-9603; trimmed enzyme; substrate specificity:  
heat resistance; acid region; production; ss.

XX OS Bacillus sp.

XX FH Key Location/Qualifiers  
FT CDS 483..3338

FT FT /\*tag- a  
FT /\*product= Trimmed\_enzyme  
FT 483..569  
FT /\*tag- b  
FT mat\_peptide 570..3335  
FT /\*tag- c

XX PN JP09271385-A.

XX PD 21-OCT-1997.

XX PF 01-APR-1996; 96JP-0134492.

XX PR 01-APR-1996; 96JP-0134492.

XX PA (AMANO) AMANO PHARM KK.

XX DR WPI; 1998-003022/01.

XX DR P-PSDB; W37372.

XX PT DNA encoding a trimmed enzyme - used in the production of a trimmed  
XX enzyme with wide substrate specificity

PS Claim 3; Page 10-13; 15pp; Japanese.

XX The present sequence encodes a trimmed enzyme isolated from  
CC Bacillus sp. APC-9603. The present specification also describes:  
CC (1) a vector containing the DNA encoding a trimmed enzyme; (2) a  
CC recombinant microbe transformed by the DNA; and (3) preparation of  
CC a trimmed enzyme in which the microbe of (2) is cultured and the  
CC trimmed enzyme is collected from the culture. The method can prepare  
CC a trimmed enzyme of wide substrate specificity having heat resistance  
CC and being active in acid region efficiently.

SQ Sequence 3934 BP; 1200 A; 756 C; 908 G; 1070 T; 0 other;

Query Match 65.3%; Score 1823.6; DB 19; Length 3934;  
Best Local Similarity 79.9%; Pred. No. 0;

Matches 2208; Conservative 0; Mismatches 544; Indels 11; Gaps 5;

QY 1 gattggaaacagcaacagatctgtccactattttgctctgtgtgattacacct 60  
DB 570 gacggagacagcaacaaacgtcatttcttatttttgcgcaggtgtgattacacatca 629  
QY 61 tggagttctatgtatgtgcccacaaagacggaggttgggctggaatattacacacg 120  
DB 630 tggagttctatgtatgtgcccacaaagacggaggttgggctggaatattacacacg 686  
QY 121 gttgactcttcttggagctgttgcagctgtgatatctccagaaacacagtcagtagga 180  
DB 687 acagattcgtatggggaaatttgcacatgttctgcgtccgggagttccagtaagtttga 746  
QY 181 attatcgttcgactcaatgaattggacaaagatgtgagcgttcgacgtacatagatta 240  
DB 747 attatcgttcgacccacagatgtggcaaaagatgtgacagaccgtacatagattctt 806  
QY 241 agcaaaaggaatgaggtgtggtcttgtagaagaaagcaaatgtttatattgaaaaa 300  
DB 807 agcaaaaggaatgaggtgtggtcttgtagaagaaagcaaatgtttatattgaaaaa 866  
QY 301 gatgtcgaagatgcagcgttaacccgcgtgtgaagcaaacgtattttagatgttcaaacag 360  
DB 867 gatgtcgaagatgcagcgttaacccgcgtgtgtgagcaaatgtattttagatgttcaaacag 926  
QY 361 gttcgtgttaaaccttagccagcgttaactcttggggaagmnaagcggtttacgtt 420  
DB 927 gttcgtgttaaaccttagccagcgttaactcttggggaagmnaagcggtttacgtt 986  
QY 421 catgacgacacgcaataagatatccagtgacatctcgtgaagatgacagttctgt 480  
DB 987 catgacgacacgcaataagatatccagtgacatctcgtgaagatgacagttctgt 1046  
QY 481 caagatgtaaccgctgttttggcaaggtaccttccacacatatgttggagttccagttg 540  
DB 1047 gaggatgttaaccgctgttttggcaaggtaccttccacacatatgttggagttccagttg 1106  
QY 541 gcaactgataatcaacagcttatttaaaaggtgactaacatctctacatctca 600  
DB 1107 gcaactgataatcaacagcttatttaaaaggtgactaacatctctacatctca 1166  
QY 601 ggaagatcttcctggaagaaactaacatataagtgctttaaataatgctggaataat 660  
DB 1167 ggtgagttgcccagggtgggttaccataataagtgctttaaataatgctggaataat 1226  
QY 661 ccgagttaccatctgcaacaataatttaacagcttccgcggcggttcacacgctcact 720  
DB 1227 ---agctaccatctgcaacaataatttaacagcttccgcggcggttcacacgctcact 1283  
QY 721 ttctgtaattcgttcacatcagctatgacacataataataatcctaagtcgagt 780  
DB 1284 ttctctatgctccgctccacccatcgttatalagacagattaataatccagagagtaat 1343  
QY 781 ttacaagatga---aagcgggtttaaacagatctgtgaaggttactctaagggagaat 837  
DB 1344 ttactttagatggtgagcgttatcaaaacagatcttagtaacggtgtcttgggaagaaac 1403

QY	838	cgaagcagtcgagcaatactctcgtgccatccaacagatggtcaacagcgaaagcagcgtgata	897
Db	1404	ccgaagatctctgcgcataccctctctacccaacagacgaogtctttaagaacagagaggtcatcc	1463
QY	898	cgctcgtaagtcgtctaatttcaacagtgactcattcttcagggagatgagatcttgggaataacc	957
Db	1464	ccctcgaaatgtgctggaactctctctccatattatattatctcggagagagatcctcggaacacc	1533
QY	958	tataccagaaagcgcaaaacacctttaagctctggccaccacactctcactcaagataatgct	1017
Db	1524	taaccagaagaaagacaaacacctttaaagtcgtggcgtccacactctcactcaagtcgaatgct	1583
QY	1018	ctctctcttaagcagtcgtaaaagggcttcgtatgaacaaaatccgtatacctatgacgcctgggc	1077
Db	1564	ctgcctcttaatacaaaagagcgggtgtgctcttaacaagaagctgtctccctaaagaggtccagga	1643
QY	1078	caatgctcgtcgagagacaaagctctaaaccaaaaccccttgaaaactctgtaatacactatagag	1137
Db	1644	catggcgtctgctgacgtgagaccgtctccacaaaacccctagaagaattctgtatctatgtatgag	1703
QY	1138	gtaaacagggccaaagcctctacccgaacgctgtgtgataccttatgcaactcgagatctgaca	1197
Db	1704	gtaacctctgtccaaagcctctcaaccgaacaaagcgtctgagccctctatgccaacgcgaatcgaca	1763
QY	1198	aabtgaaacagaaagacagatctgtgagccctcggtccaaabaagatccctgcgtcgagtcgaacgct	1257
Db	1764	aatgtgaacagagagcagatcagctgtgtctgacctagccgaabaacaaacccgacactgctgtgaabagac	1823
QY	1258	gataaaacatatagccccaagaaatalagaaagatgagatcatctatgaaatagtgatctgcgt	1317
Db	1824	gacaaacataatgacaacccgaagaaacataagagatgtaagatattatatagaaatgcatatgcgcga	1883
QY	1318	gaacttcccatctgagcccttaattctgggtatagaagaaataaaagggaaatatttggcctctaca	1377
Db	1864	gaacttcccatctgagcttccaactccaagatcagacaataaaagggaaatgaaatctgaccccttact	1943
QY	1378	gaabaabagaaacaaagggccctctgacaacagctanaagacgaggatctgaaatctcttaaaacaact	1437
Db	1944	gaaabaagagaaacaaagggccctctgaaabaatgtcaabaacagcgtgtgatactaatctgaagcagctt	2003
QY	1438	ggagatctactcaatgtctcagctctatagctcgtcttccgcactcaacagctgctcgatgaaacatgata	1497
Db	2004	ggacttccccaatgtctcagctctcagctcgtctcttccgccttataacagcgtctgatagaaagcgaat	2063
QY	1498	ccaaaccccaagataactctgaggtctatgagccctcgagaaactatgaaactgtcccgaaagggcagctat	1557
Db	2064	cccaaccccaatataactctgaggtcctatgactcctcgttaactataatgcttcggaaagggcagctat	2123
QY	1558	gctaaacaaatcggaaatggtaaatgtctcgtatataaaagagctttaagaaatggtctcttcaatc	1617
Db	2124	gcaaacgagatctgcaaaacgagcaacactcggatctaaagagctttaagaaatggtctctctccctc	2183
QY	1618	catctcgtgaacacatctggggtctaaacatgagatgtctgtctataatcatcaactcttgacaagaa	1677
Db	2164	catctcgaacacacatctgagatccaacaatgagatgggtgatataaaatcaataactcttgccacaaca	2243
QY	1678	atctctgactctgataaagaattgtatacagaagaatactataccggtac--gatgatccaggttaat	1736
Db	2244	atactcgtactctgtataagatctgtctcccgcaataattatataccggaacagatgatacgtcgagac	2303
QY	1737	tatctccaaacgatacaggtctactgagaaatgaaatgtacgacgngngaaagggccaaatggtctcaaaa	1796
Db	2304	tacacccaaatgggtctaaagctaaagaaagaaagcgtgaca--gctgaaagcggccaaatggtctcaaaa	2362
QY	1797	attataatctatccctctaaagatctgggtcacaatgagatactatactatgagccgtctccgtt	1856
Db	2363	attataatctatgactcactctaaagatctgggtgagaaatgagatcaatatactgagccgtctccggtt	2422
QY	1857	tgaacttaatgagcgtcgtctctgaaagaacagatgctccaagcgtcgtccctcgagcgtcatgac	1916
Db	2423	tgaacttaatgagcattactcttgaaabaatatacaaatgagaaagacgagacaaaggtctcatgac	2482

[illegible]





QY 1585 ataaaggttaaggaatggtttcttcaactcactcgttgaaacacatggtggttaacatg 1644  
 DB 1029 atcccgaggttaaacgctgacgttcatagcttgcaatcaacgacgtacgtcattg 1088  
 QY 1645 gatgttgctatcatcactttgacacgaatctgactgtatgaataatgttaca 1704  
 DB 1089 gatgttgctatcatcactttgacacgaatctgactgtatgaataatgttaca 1704  
 QY 1705 gaataatcatcgt-taacatgatccagtaataatlaacaaacgacgtgactgaat 1763  
 DB 1149 ggtgtattcttccgttgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1208  
 QY 1764 gaatttcgacgaaaggaacgaatggttcaaaaatlaattatgatttccctaaatg 1823  
 DB 1209 gatattgc-ttccgaaacgaataatggttggaataatgattatgactcgttgcattg 1267  
 QY 1824 ggtcaatgagatatacatatgacgttccgttttgaacttaagggcgtcgttggaaga 1883  
 DB 1268 ggtcgaagaatatacatgtaacacggaattcgttgaattcaatggaatttgaattga 1327  
 QY 1884 cagatgtccaaagctcgtcgcgaacgttcaatgtaataacgaagaatgtcacttaacg 1943  
 DB 1328 aacgatgaaagcgtgtcgcgaatgttgatacactcgaacccgttctattctgtattcg 1387  
 QY 1944 tgaacatgagcgggtggaacactcgtcactcgaatgatacagcttctgacaaagagc 2003  
 DB 1388 agaaagatgagcgtgtcgcgaacacgtcgttgcgttgaacaaagacgagtcaaacgc 1447  
 QY 2004 tcaaaaagcagatggaagatgagcgtgttgaatgaacaaatgaacgaacgttggaagcga 2063  
 DB 1448 agaaagatgagcgtgtcgcgaacacgttgaacaaatgaacgaacgttggaagcga 1507  
 QY 2064 tgccttgaattcctcgtcgcgaacgttgaacgaacgttggaacgaacgttggaacga 2123  
 DB 1508 cagcttgaacgtcgcgaacgttggaacgttggaacgttggaacgttggaacgttgga 1567  
 QY 2124 taagaatgagcgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 2174  
 DB 1568 gaaatgaacgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 1627  
 QY 2175 gaaatgaacgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 2234  
 DB 1627 gaaatgaacgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 1687  
 QY 2234 caatcctaattcgtcgcgaacgttggaacgttggaacgttggaacgttggaacgttgga 2294  
 DB 1687 taatgacatgaaatgaaacgaacgaacgaacgaacgaacgaacgaacgaacgaacga 1747  
 QY 2295 tatgaactcaacgaacgttggaacgttggaacgttggaacgttggaacgttggaacgt 2354  
 DB 1748 gctgttcccaaggaatcgttcttgcatacgacgaacgaacgaacgaacgaacgaacga 1807  
 QY 2355 cggcaacgaacgaacgttggaacgttggaacgttggaacgttggaacgttggaacgt 2414  
 DB 1807 aggaatgaacgaacgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 1867  
 QY 2415 agctcaatcatcagatgttctcaactattatagcggaacgaacgaacgaacgaacga 2474  
 DB 1868 gagcgaacgaacgttggaacgttggaacgttggaacgttggaacgttggaacgttgga 1927  
 QY 2475 ccgaacgttcgacatgacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacga 2528  
 DB 1928 ccgaacgttcgacatgacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacga 1981

RESULT 7  
 ID V36920 standard; DNA; 2319 BP.  
 AC V36920;  
 XX  
 XX  
 DT 21-DEC-1998 (first entry)  
 XX

DE Thermotoga maritima 6GP3 pullulanase gene coding region.  
 XX Glycosidase: 6GP3; thermostable enzyme; oligosaccharide; glucose;  
 KM sugar; baking; textile; detergent; pullulanase; ss.  
 XX  
 OS Thermotoga maritima strain 6GP3.  
 PN  
 XX W09824799-A1.  
 XX  
 PD 11-JUN-1998.  
 XX  
 XX 08-DEC-1997; 97WO-US22623.  
 PF  
 XX 10-OCT-1997; 97US-0949026.  
 PR  
 PR 06-DEC-1996; 96US-0056916.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 PI  
 PI Bylina EJ, Lam DE, Mathur EJ, Swanson RV;  
 DR  
 DR WPI: 1998-362407/31.  
 XX P-PSDB: W49871.  
 PT Glycosidase enzymes from organisms of the genera Staphylothermus,  
 PT Pyrococcus and Thermococcus - for deriving sugar from  
 PT oligosaccharides, useful in the e.g. food processing, textile or  
 PT baking industries  
 XX  
 PS Claim 1: Fig 14a-e; 92pp: English.  
 XX  
 CC This isolated polynucleotide comprises a coding region for  
 CC pullulanase 6GP3 (see W49871) from a Thermotoga maritima clone  
 CC (6GP3) that grows optimally at 85 degC. The sequence shows 53%  
 CC nucleic acid identity to an alpha dextran 6 glucanohydrolase of  
 CC Caldocolum saccharolyticum. The invention provides 18  
 CC polynucleotides (see V36907-24) coding for thermostable glycosidases  
 CC (see W49858-75) having glucosidase, alpha-galactosidase,  
 CC beta-galactosidase, beta-mannosidase, alpha-mannanase, endoglucanase  
 CC or pullulanase activity. Vectors and host cells are also claimed.  
 CC A method is provided for producing the enzymes by recombinant  
 CC techniques. A claimed method for generating glucose from soluble  
 CC cell oligosaccharides comprises contacting a sample (selected from  
 CC dairy products, fruit juice, detergent, textile, guar gum, animal  
 CC feed, plant biomass or waste product) containing oligosaccharides  
 CC (selected from maltose, cellobiose, lactose, sucrose, raffinose,  
 CC starchose, verbascone, cellulose, starch, amylose, glycogen,  
 CC disaccharides, polysaccharides and pullulan) with one of the  
 CC claimed glycosidases such that glucose is produced.  
 XX  
 SO Sequence 2319 BP; 760 A; 512 C; 584 G; 463 T; 0 other;  
 Query Match 9.9%; Score 276; DB 19; Length 2319;  
 Best Local Similarity 51.3%; Pred. No. Be-70;  
 Matches 933; Conservative 0; Mismatches 817; Indels 68; Gaps 10;  
 QY 926 actactattcagagatgactgttgggaatacctatcacacgaagaacgaacatttaag 985  
 DB 440 actactattcagagatgactgttgggaatacctatcacacgaagaacgaacatttaag 499  
 QY 986 tctggtcacaacattctactcaagaatgcttctcttataatgacgtgcaacggttctg 1045  
 DB 500 tctggtcaccctgttcttaagtggttaaaagtgcttcttcaaaaacggaagaacacg 559  
 QY 1046 taacaaaatcgtactatgacggcat---cgggacatggtgtgtggaagaacaggtta 1102  
 DB 560 aacggtacacaggtgtggaacatgatacaaggaacggtgtgtggaagcgttctg 619  
 QY 1103 atcaaaacgttgaataatgttgatataatgagtgaaagcgaacggtctacccgaa 1162  
 DB 620 aaggtcgtctcgaagagtgatcttccctcattcagcttgaacgaacgaacgaacga 679  
 QY 1163 cgtgtgtatccttatgcaactgcatgacacaaatgaaacgagagcatgattctg 1222

Db 680 caacccgctcgaatcccttctacgaagcggttcaagcaaaacacccaagagagccgctgtgga 739  
 Qy 1223 accgtgtcaaaacagatccgtcgtcgtggaacagtgataacatatctacgccaagata 1282  
 Db 740 atctctgcagagcaaaacccgaagagatgaggaataacagacaggggacgcaaaatcgaaagat 799  
 Qy 1283 taagaaatgaggtcatctatgaaatgagtgctcgtgacatttccattgacctaattcgg 1342  
 Db 800 acgaaagcgcgcatatctatgaaatlaacacaaagcatalcaacagatcgaataaccgcg 859  
 Qy 1343 gttatgaaataaaggaagatatttgctcttccagaaaaagaaagcaaaagggccctgaca 1402  
 Db 860 ggggaaataaacaagagccctcctacgtcgggctcaccgagaagaaacagaaagagccggcg 919  
 Qy 1403 acgtaaagagcgggagatgatctcttcaaaaacaaactlggatctacatgcagcttctgc 1462  
 Db 920 gtgtgaaataacagccttctgcacctctgtgaaactcgggtgttaccacagcttcatatttc 979  
 Qy 1463 cgtgttctgcatacgaagctgcagtgaaactgatacca-----accgaagataatt 1513  
 Db 980 ctctctcttgatttctcaacagcgacgaactcgaataaagatttctgagaagtgatacact 1039  
 Qy 1514 ggggtttctgacccctgcgaactatgattgttctctgaaagcgagatgatacaaatgagatg 1573  
 Db 1040 ggggttctgacccctcctacgtctcaatggtctccggagggcagatctcaacccgataccaaa 1099  
 Qy 1574 g---taatgctgataaagagagttlaaagaaatggtctcttccatccatcgltgaacaca 1630  
 Db 1100 acccaacacagagatcatgagaaatgcaaaagaaatggtcacaagccctcacaacacaggtta 1159  
 Qy 1631 ttgggtgttaacatgagatggtctctataatcaatacacttctgcacgc---aaatctctgact 1667  
 Db 1160 taagtgtgatatatgaaatgacatggttccctcacaacccctacgggtatagcgaaactcttcggt 1219  
 Qy 1688 tcgataaataatgtatcagaataatattatataccgcta--cgaatctcaggtatataccaaag 1746  
 Db 1220 tcgatacgaaggtgcgttactactctcaagaatcgaacagagcggtgcctatttgaaag 1279  
 Qy 1747 gatacaggtacgtgaaatgaaatltgcangcnaagaaagccaaatggttcaaaaaattatatt 1806  
 Db 1280 aaagcggtatgtgaaacgtctacgcgaagc--gaaagaccatgataagaaattatattatgctc 1338  
 Qy 1807 gattcccttaagttatgtgtcacaatgatatcatatgacggtccgttcttgacttaagt 1866  
 Db 1339 gataccgtccactctctgtgaaagagatcatcaatcagagattcgaagttctgatalcagaatg 1398  
 Qy 1867 ggcgtgtcttgaaagagacagatgtccaaagcgtccctcgagagctctcatgctatattca 1926  
 Db 1399 ggtctctctcabaacaaagacatgctcgaagtcgaagagagctctctataaattcgatcca 1458  
 Qy 1927 ggaattgcaacttaccggtgagcgaatggaacggtlgaacacttgcacatgcagatgacag 1986  
 Db 1459 acatcatctctcagcggaacggtggtgtgagtggagacac-----gatacag 1508  
 Qy 1987 ctctctgcanaaagagatcacaagagcatgagagtaagcgtgtttaaatagacaatttca 2046  
 Db 1509 gtttggaaagagcgaatgctgcgc--ggaacacacgtlgaacacttcaacagatgattcaga 1566  
 Qy 2047 aaacgctgtgagcgaatgtcttctgattcttcgcgtcacaaggttttgcagacgtgcacaca 2106  
 Db 1567 gaacgaataaaggggtctcgttcaacccgagcgtcaagagattcgtcatagggagatcac 1666  
 Qy 2107 ggtctaaacgtatgcaattaaagatlgcgtltgaaggagattttaaagc----- 2154  
 Db 1627 ggaagagaaacaaagatcaaaaggggtgtgtgtgaaagcataaactacgacggaacatc 1686  
 Qy 2155 -----ttaccttctcaccaggtgagacaataatactatgtcacaagttcatgataac 2205  
 Db 1687 atcaaaagttctgccttgcgtcagaaagaaactataaactacgacaggtgtcacagacaac 1746  
 Qy 2206 taacaccttggagcaaaatagccctaagcaatcctaattgatt-----cc 2250

Db 1747 cacacactgtgagcaagaactacacttgcgcgcgaagctgataaagaaagaaatgagcc 1806  
 Qy 2251 gaacggtatcggtatataaalgatgaactcgcacaaagcagttgtatgacctcacaagc 2310  
 Db 1807 gaagaaataactgaaaaaacgcgccgaagaaactgtgtgtcgatcttcttccaaaggt 1866  
 Qy 2311 gtccatctatgcgaagcggggaagaaatgctctgtanaaaagcggaacagacaatagt 2370  
 Db 1867 gtctcttctcccccacggaagcgaggaactctgcagagcagcaattcacaacgaactcc 1926  
 Qy 2371 tataatcagagcgaatgcgtcaatagatgttgattgagcgaggaaagctatataccgat 2430  
 Db 1927 tacaacgcccctatctcgaataaacggtctgatatcaagaagaaactcctcgttcatgac 1986  
 Qy 2431 gtttcaactatlaaagcgggtcaatccacacttgcgttcttgataccacagcttccgatg 2490  
 Db 1987 gtgttcaatttaccacaaggggtctcatataaactcgaagaaagaaacccctgttcttaagctg 2046  
 Qy 2491 acgacagctaatgaaatcaatlaagccacctccaaattcctaataatgtccagagaaacagtg 2550  
 Db 2047 aaaaacgtgaaagagatacaaaaacacacttgaaatttctcccggtcggtgagaagatagtt 2106  
 Qy 2551 gccatgaaatlaactgtatcgtatgataaagacaagaggggaaatatacatgtgtttat 2610  
 Db 2107 gcgttcatgcttaagagacacagcaggtgtgtgattccctggaaagacatcgtgtgattac 2166  
 Qy 2611 aaccacaataaanaactgtaagacacatcaatttgcgcgagcggaatggtgcaatcaatgct 2670  
 Db 2167 aatgaaataacttaagaaagacacatacaaaactgcgcagaaagaaatgaaatgaaatggtgtg 2226  
 Qy 2671 acgagcgttaaggttaga 2688  
 Db 2227 aacagccagaagagcgga 2244

RESULT 8  
 T93691  
 ID T93691 standard; DNA; 2319 BP.  
 AC T93691;  
 XX  
 DT 12-MAR-1998 (first entry)  
 XX  
 DE Thermotoga maritima pullulanase encoding DNA.  
 XX  
 KW glycosidase; thermostable; textile; food processing; pharmaceutical;  
 KM detergent; baking; industry; Thermococcus; Staphylothermus;  
 KM Pyrococcus; glucose; soluble oligosaccharide; pullulanase; ss.  
 XX  
 OS Thermotoga maritima.  
 XX  
 Key Location/Qualifiers  
 FH 1..2319  
 CDS /\*tag- a  
 FT /product- Pullulanase  
 FT  
 XX WO9725417-A1.  
 XX  
 PD 17-JUL-1997.  
 XX  
 PE 10-JAN-1997; 97WO-US00092.  
 XX  
 PR 13-SEP-1996; 96US-0712612.  
 PR 11-JAN-1996; 96US-0583787.  
 XX  
 PA (RECO-) RECOMBINANT BIOCATALYSIS INC.  
 XX  
 PI Byline EJ, Lam DE, Mathur EJ, Swanson RV;  
 DR WPI; 1997-372858/34.  
 DR P-PSDB; W34567.  
 XX  
 PT New thermostable glycosidase(s) - from Thermococcus, Staphylothermus



PT and Pyrococcus, used in the textile, food processing,  
PT pharmaceutical, detergent and baking industries

PS Claim 4: Fig 14: 82pp: English.

XX The present sequence encodes pullulanase isolated from Thermotoga  
CC maltima. The enzyme or its encoding nucleic acid sequence is used for  
CC generating glucose from soluble oligosaccharides. The enzyme can be used  
CC in the food processing, pharmaceutical, textile, detergent and baking  
CC industries. The enzyme is also used to treat lactose intolerance, as a  
CC diagnostic reporter molecule, in corn wet milling or in the fruit juice  
CC industry. The enzymes can be used to hydrolyse guar gum to remove  
CC non-reducing terminal mannose residues. The nucleic acids encoding the  
CC enzyme may be used to generate probes to identify similar sequences.

SO Sequence 2319 BP; 760 A; 519 C; 577 G; 463 T; 0 other;

Query Match 9.7%; Score 271.2; DB 18; Length 2319;  
Best Local Similarity 51.2%; Pred. No. 2e-68;  
Matches 930; Conservative 0; Mismatches 820; Indels 68; Gaps 10;

QY 926 actattactcaagaagatgattcttggaataactatacaagaagaacaacatttaag 985  
DB 440 actactactcaagaagatgattcttggaataactatactcagaagaagatattcagaag 499  
QY 986 tctgggcaacaactctactcaagaataatgtctcttctatgacagtgcaagggcttcg 1045  
DB 500 tctgtcccccgttcttaagtgagtgatgagtgctctctctcaaaaacggaagaacag 559  
QY 1046 taacaanaatcgtacctatgacgcat---cgggcatagtgtgttggaagcaacggtta 1102  
DB 560 aacgctacaggttgtgaacatgatacaagaagaacgggtcttggaagcggtgtgtg 619  
QY 1103 atcaaaacttgaataatgtgattacatgataagtaacagcgcaagcttaccgga 1162  
DB 620 aagcgatctcgaagtgatgtctctctatcaagcttgaataacatgaagaatatacaa 679  
QY 1163 cggcgttgatctctatgacactgcgactgtcacaagaatgaacgaagacatgattgag 1222  
DB 680 caaccgttcgactcttctatcgaaagcggtttaacgaacaacacaagaagagcgctgtgta 739  
QY 1223 acctgtgtaaaacagatcctgtgtgacagatgataacataatcgcgaagaata 1282  
DB 740 atctgtccaggaacaacccaagaagatggaacacgacgggacccaataatcgaaagt 799  
QY 1283 tagaagaatgaggtcatctatgaaatgagatgtcgcgagacttccattgcccctaattcg 1342  
DB 800 acgaagacgcgataactatgataacatacagatgcgacatcagaagactcgaaaaactcg 859  
QY 1343 gtaataaataaagaagatatttggtcctctacagaanaaagaacaagggccctgaca 1402  
DB 860 gggtaaaaaaacaagccctctatctcgggtcaccggaagaagaagaacggcgag 919  
QY 1403 agctaaagacgggagatagattccttaaaacaacttgggattctcatgttcagttatgc 1462  
DB 920 ggtgtacacaagaagccttcgcacactgtgtgaactcgtgttacaacgcttcatattc 979  
QY 1463 atgttttcgactcaacagtgctcgaatgaaactgacca-----accagaatatt 1513  
DB 980 cttctcttgattctcaacaagcgcaacaactcgataaagattctgagaagatatacaact 1039  
QY 1514 ggggttatgacctcgcaactatgattgttcctgaagggagatagtaacaaatgc---ga 1570  
DB 1040 ggggttatgacctctacactgttcattgttcggagcgagatatacaacgactcccaaaa 1099  
QY 1571 atgtgaatgtctgtataaagaagtttaagaaatgttcttcaactccatcgtgaacaca 1630  
DB 1100 acccaacacggaatcagaagaagtaagaagaatgttcaaaaccccttcaacaacagtgta 1159  
QY 1631 ttgggttaacaatgattgtctataatcaatcacttgcacgc---aatcttgact 1687  
DB 1160 taggtgtgattatgataatgattgttccctcacacactaaggtatagcgaaactcttcggt 1219

QY 1688 tcgaataaaatgtaccagaataattataccgta-cgatgatccaggttaattataccaag 1746  
DB 1220 tcgatacagcggtgcgactactctctacagatcgacaagaacaggtgctatttgaagc 1279  
QY 1747 gatcaggtctgtgaatgaaatgtcagcngcgaaggaagcgaatgttcaaaaattattat 1806  
DB 1280 aagcgatgtgtgaatgtcagcagcagc-gaaagcccatgtgaaatattatcagtc 1338  
QY 1807 gattcccttaagttatggttcgaatgatactatgagaggttcggttcttgaataag 1866  
DB 1339 gatccgctcaactactgtggaagagatatacacaagcagatcaggttctgacagag 1398  
QY 1867 ggcgtcttggaagaagacagatgttccaaagctcctcgagcttccatctataatca 1926  
DB 1399 ggtctatcagacaanaagcaatgtcgcgaagtcgaagaagccttctataatcgtacca 1458  
QY 1927 ggaattgcacttaagtgatgacgaatgagcgtgtgaaacctctgcactgcagaatgacag 1986  
DB 1459 actatcattctctacgcgcaacccgtgtgtgtgagtgaggacac-----gatacg 1508  
QY 1987 ctcttcgacaagaagcctcaaaaagcagtgaggatgaggtgtttaaagaaatttaca 2046  
DB 1509 gtttggaagaagcagatgtgcgc---ggcacacaagctgagcagcttcaacagatgtcaga 1566  
QY 2047 aacgcgttggaagcaatgtcttgaattctccgcctcaaggttttcgacaggtgacaca 2105  
DB 1567 gacgcaataaagggttcgttcgttcaacccgagtgcaaggtatctgtcgttggaagatgc 1626  
QY 2107 ggttaactgtatgacaattagaatgaggtgtgagggagattaaatgac----- 2154  
DB 1627 ggaagaagaacaagaatcaaaaggggtgtgtgtggaagaataaactcagacgaanaactc 1686  
QY 2155 -----ttaaccttcacaggttgagacataatgattatgataagtaagtaac 2205  
DB 1687 atcaaaagtttcgccttgcattcacaagaagaataactaactacagcagctgtcagcaaac 1746  
QY 2206 tacacctttgggacacaanaatagccccaagcaatcccaatgat-----cc 2250  
DB 1747 cacacactgttggaacagaactacactcttcgcgcgaagaagcgtataagaanaagatgac 1806  
QY 2251 gaagcgatcgtgaattaaatgtgaactcgcgaacagcaggtgttatgacctcacaagc 2310  
DB 1807 gaagaagaactgaaanaacgcccgagaactgtggtgcataactctctcctcaaggt 1866  
QY 2311 gtccacatcatgaaagcggggaagaagaatgcttcgtanaaagaacggcgcaagcaatagt 2370  
DB 1867 gtcccttcctccacaggaagggagagactcctgcaggaacgaatctcaacgacaactcc 1926  
QY 2371 tataatgcagcgatgcgttcaatgagttgtatgtgagcaggaagctcaatctcaga 2430  
DB 1927 tacaacgcctctatctcgtataacgcgcttcgatataagaagaataactcagttcatagac 1986  
QY 2431 gtttcaactatataagcggtcactcaacactcgtctgataacccaagccttcgcagatg 2490  
DB 1987 ggttcaatattcaacaagaaggttctcatalaaactgaagaagaacacactgcttcaagcgt 2046  
QY 2491 acgacagcctaataatgaaatagccacactcccaacttccaaatagtcacaagaacagtg 2550  
DB 2047 aaaaagcgtgaagagatcaaaaacacactcgtgaattcttcctccgggctcggaagaatagtc 2106  
QY 2551 gcctatgaatataatcgtatgataatgaagaacaaatggggaataatcattgtgttatt 2610  
DB 2107 ggcgtatgtttaaagaacacgaggtgtgtgattcccttggaacacacatcgtgtgatttac 2166  
QY 2611 aaccacaataaactgtatgacaacatcaatttccgagcggtggaatggaacatcaatgct 2670  
DB 2167 aatggaactttagaagaacacatacaaaacttgcagaaagaaatggaatgtgtgtgtg 2226  
QY 2671 acgagcgtlaaagtagga 2688  
DB 2227 aacagcgaagaagcgga 2244

```
RESULT      9
X90036
ID   X90036 standard; DNA; 2601 BP.
XX
AC   X90036;
XX
DT   17-SEP-1999 (first entry)
XX
DE   Ferriodobacterium pennavorans Ven 5 thermostable pullulanase DNA.
XX
KM   Ferriodobacterium pennavorans; Ven 5; thermostable pullulanase; pullA;
XX   Industrial saccharification; ss.
XX
OS   Ferriodobacterium pennavorans.
XX   MO9935274-AZ.
XX   PN
XX   PD   15-JUL-1999.
XX   PF   05-JAN-1999; 99MO-1B00069.
XX   PR   07-JAN-1998; 98US-0003834.
XX   PA   (NOVO ) NOVO-NORDISK AS.
XX   PI   Duffner F, Jorgensen PL;
XX   DR   WPI: 1999-430401/36.
XX   DR   P-PSDB; Y24380.
XX
PT   Nucleic acid encoding Ferriodobacterium sp. Ven 5 thermostable
XX   pullulanase, useful for Industrial saccharification processes
XX
PS   C1a1m 1; Page 40-41; 52pp; English.
XX
CC   The present sequence encodes Ferriodobacterium pennavorans Ven 5
XX   thermostable pullulanase. Thermostable pullulanases are useful in
XX   CC   Industrial saccharification processes (see EP03909).
XX
SQ   Sequence 2601 BP; 881 A; 494 C; 596 G; 630 T; 0 other;

Query Match      9.64; Score 269.6; DB 20; Length 2601;
Best Local Similarity 50.7%; Pred. No. 6.2e-68;
Matches 953; Conservative 0; Mismatches 856; Indels 71; Gaps 10;

OY   931 tatcagagagatgactcttggaatacctatacacagaaagcaacacctaagaatctgg 990
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   733 tatcagatggtccacatcggaattggaatacaccctgaaagacaagatagaggtttgg 792

OY   991 gcacaaactctactcaagaatgtctctctcttaagacagtgcaaggggtctgttaca 1050
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   793 tcaccgtttcaaaaactgtgatgtactacttcaaaaatgtggaatgataaagaacca 852

OY   1051 a---aaatcgtactatgaagcgcatcgggcattgtgtfvgggaagcaacggttaacca 1107
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   853 aagcaaatgtgttccaaatgaatatalatagcaagggacatgaggaagcgctcttgaaag 912

OY   1108 aacctgaataatgtgtaatacatgtatgagtaaacagggcgaagccttaaccgaagcgt 1167
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   913 aactgggaatgtatgtctctcaaaaatcaagttactcagctatggaatatagaatca 972

OY   1168 gtatccttcaactgaactgcatgtacacaaatlgaaacgaagagcatgattgttgacctg 1227
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   973 gtgatattactctcaaaaagcgtaacgaagatctctgcaagagtggtatcatgtacttc 1032

OY   1228 gctaaacagatccctgtctgtctggaacagatgataaacatatattacgcgaagaataga 1287
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1033 aagcaaaaacaaatccagagagctgtggaaaaagtcgcagaccaccactagttgcccacag 1092

OY   1288 gatgaggtcatctctatgaatgagatgctcgtgacttccattcgaatgacctaatgtggcatg 1347
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB   1093 gatgcgataataatcacgaataccacatcgcagatatagcacggactcgataattctgctgc 1152

OY   1348 aaaaataaagggaaglaatttgcctcttacagaaaaaagaaagagccctgcgaacgta 1407
    || | || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1153 aagacaagaagcacttacttctgtggacttcacgaaaaaagcacaagaggtctctaattgttga 1212

OY   1408 aagacgggataagattcctctaaacaaacttggatatactatgtcttaagtctgtc 1467
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1213 acaactgtcttgacccaccctgtgaactcggggttactacgttcatattcttccaatg 1272

OY   1468 ttgcgactcaacagtgctcgatgtgaactgtaaccccaacaaat-----aatggaggt 1518
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1273 ttgatttctcggaacagcgatgaagctgataaagaatttgaagaagttcaacttgggt 1332

OY   1519 tatgacctgcgaactatgatgtctctgaagggagatgtatacgaatgcg---aaggtc 1575
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1333 tatgaccataactccttcttaacagttccagaaggaagccttaccacagaccacatcaacc 1392

OY   1576 aatgctgtataaagaagtttaagaaatgtctcttccactcactgcgtgaacacattggg 1635
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1393 tacacaagaaattatcgaagtcacaaacaaatgtttaaagcacttccatgaaaatggtataag 1452

OY   1636 gttaacatgagtgtgtctataatcacactttgc---acgcaaatctctgactcgat 1692
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1453 gtatccctgtatagtcgtctccacatacagggcgttggttgaatgtccatccgac 1512

OY   1693 aaaaatgtaacagaaatataatlaaccgta-cgatgatccaggttaataacaaacgata 1751
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1513 cagcgagttccatattactcttaccagaaattgcacaaacaggttgatatttgaacgaagtc 1572

OY   1752 ggtacttggaaatgaattgcgaacngcngaaagccaaatgtctcaaaatattatattgttc 1811
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1573 ggtctgtgaatgtatgtatgcagacg-gaaagacggaatgataaataacatacgtgcacac 1631

OY   1812 ccttaaglatgtggtcgaatagatcatatgaacggttcogtttgaacttaagtcgct 1871
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1632 actcaaatgtgtgttcacggaatacaaatgtatgttccagatttgaacgaatggattc 1691

OY   1872 gcttggaaaagacagatgltccaaagctgcctcggaactcatatgaatcaacgaagat 1931
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1692 aatgataaagltgacaaatgtctgcgataaaaatcgaactcgaagatagagctctgtc 1751

OY   1932 tgcacttaaggtgtagacacaggaaggttggaactctgcacatgcgaatgatcgccttc 1991
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1752 ggtcttcaacgtgaaacctgcggcgttctgggagcccaattgaattg----- 1801

OY   1992 gacaaaaggagctcaaaaagacatggagatagcggttcttaatatcaaatltaacgaagac 2051
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1802 --gaaaagccgattgttgggtataggtatcgcgcgttcaacgaatgaattcgaagatgc 1859

OY   2052 gtggacggcaatgtcttctgaattctcgcgtcaaggttttcgacaggtgcgaacaggtc 2111
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1860 actcgtgtgtctgttcttccaaagcaactgtcaagggctcttcttggcgtctcttcgcaa 1919

OY   2112 aactgtatgcaatttaagaatgctgttgaagggtatgaataga----- 2153
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1920 aagaacaggtgtataaagcggaggttgcaggttagcatcgaatatacgaacgaatlcgaag 1979

OY   2154 cttaactcttcaacaggttagacaatataactatgtatgaacaagatgatataactaccct 2213
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   1880 ctttgcgaagaagaccacagaagaactatacaactaactcgaagtctatataaccacactc 2039

OY   2214 ttgggacaaaatgaccttaagcaatctcctaattcttccgaagcggatcggatlttaaat-- 2270
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   2040 gtgggacaaaactaactctgcgacagcagcgatataacgltaaagtggactggaagaat 2099

OY   2271 -----ggatgaactcgcacaagcaggtgttataactcacaagggcgttccatc 2318
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   2100 gctcaaaagacgtctcaaaaagctgtctgtgtatatttgtacatcaacaaggtataacatt 2159

OY   2319 catgcgaagcgggggaagaatgtcttgtaaaaagcggcgaacgaacgaatagttatgtc 2378
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   2160 ctgtcagtcaggttcaagaattcgcgttaagaacaaagaatctcgatgagaactcttacaacac 2219
```



Db	2699	cggcgagaaaggagacacattcttcttcttcgcgcacccgtcggtcttcgccttcgcgaaccagcg	2758
OY	2111	taactcgtacgtacatttaagaaatgctgtgtgagggagatatacgaacttacc-----t	2161
Db	2759	ggcgagagaaacagtgaaagctcgccatctgcgcgggagctctcgagcgcgcgctgtcttt	2818
OY	2162	cttcacccaggttgagaaatatactatgtctatcaacagatcatgatactacacccctctgggaca	2221
Db	2819	ggcaaccgcgtccatcaatcaatcaatcttcgtgaatgtcatgacacacatctgtcttgagata	2878
OY	2222	aaatctagccctaaagaaatctcttaattcttcggaaagcgatccgataatgaatgatactgc	2281
Db	2879	agatgtgagcgcgcccaacacatgatgtgacggaatctgcgcgcgaagaacgcggagaaagcgg	2938
OY	2282	cacaaagcagctgtatgacccctcaacaaagcgcttccatctacgtcaaaagcggggaaagaaatgc	2341
Db	2939	cgaacgcgagatcttcgtgtctgtgcgcaaaaggaaatccgctttctgacagcgccagagattct	2998
OY	2342	tccgtctaaagaaacggcgcaacgcgaacaaatagatataatgcagcgcgatgcgggtctcaatgattg	2401
Db	2999	atcgagacgaagaagcggtcgatgtggaatctaacgcgatccggatcggaatgcggtatcaatcagctgg	3058
OY	2402	attgagagcaagaaagctcacaataatccagatglttctcaactatattatagcgggactaatcacc	2461
Db	3059	atctggtagagcggaaaaagccgcataatgaagaagatccgtaactttaa---ggaattgatcccc	3115
OY	2462	tctgcgtctgtatcccaagccttcgcgataatgacagacgtcaatgaaatcaaatgacacctcc	2521
Db	3116	tctgcgcgtgtgcatactgcatcttcgcctccgcacagaaagcggaagtgtgtgtcatcttca	3175
OY	2522	aattcct 2528	
Db	3176	cgcttct 3182	

The present invention describes a computer readable medium which has the nucleotide sequences SEQ ID NO:1 to 391 (V521314 to V52524) recorded on it, or a representative fragment or a sequence at least 95% identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1 to 391 (V521314 to V52524) are genomic fragments from *Streptococcus pneumoniae*. The present invention also describes an isolated nucleic acid molecule encoding a homologue of any of the fragments of the *S. pneumoniae* genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced by a process comprising: (a) screening a genomic DNA library using as a probe a target sequence defined by any of the sequences in SEQ ID NO:1 to 391, identifying members of the library which contain sequences that hybridize to the target sequence and isolating the nucleic acid molecules from the members; or (b) isolating mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid molecules whose nucleotide sequence is homologous to amplification primers derived from the fragment of the *S. pneumoniae* genome to prime the amplification and isolating the amplified sequences. The computer readable medium can be used in a computer-based system for identifying fragments of the *S. pneumoniae* genome of commercial importance, or expression modulating fragments of the *S. pneumoniae* genome. Products from the present invention can be used in diagnosis kits and assays, and pharmaceutical compositions and vaccines for *S. pneumoniae*.

Sequence 9541 BP; 2886 A; 2081 C; 1691 G; 2883 T; 0 other;

Query Match	6.1%	Score 171	DB 19	Length 9541
Best Local Similarity	51.0%	Pred. No. 5.7e-39		
Matches 557	Conservative 0	Mismatches 518	Indels 17	Gaps 6

RESULT 11	
V52265/C	
ID V52265 standard; DNA: 9541 BP.	
XX	
AC V52265;	
XX	
DT 23-OCT-1998 (first entry)	
XX	
DE Streptococcus pneumoniae genome fragment SEQ ID NO:132.	
XX	
KM Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;	
KM computer readable medium; vaccine; pharmaceutical composition; ds.	
XX	
OS Streptococcus pneumoniae.	
XX	
PN MO9818931-A2.	
PN	
PD 07-MAY-1998.	
XX	
PF 30-OCT-1997; 97MO-US19588.	
XX	
PR 31-OCT-1996; 96US-0029960.	
XX	
PA (HUMA-) HUMAN GENOME SCI INC.	
XX	
PI Barash SC, Choi CH, Dillon PJ, Dougherty BA, Fannon M;	
PI Kunsch CA, Rosen CA;	
XX	
DR WPI: 1998-272225/24.	
XX	
PT Computer-readable medium with recorded Streptococcus pneumoniae	
PT polynucleotide sequences - useful in diagnostic kits and assays, and	
PT pharmaceutical compositions and vaccines for Streptococcus	
XX pneumoniae	
PS	
XX	

Claim 1: Page 900-906; 1409pp; English.

OY	1277	agaataagaaatgaaggatcactctgtgaaatgaaagtgcacgttcaatttcacatggcccca	1336
Db	2004	AAATTCATGTGAAGCAGATTATCTGTGAAATGCACATTGCGTATTTGACTAANTACCTTA	1945
OY	1337	attcggtcatgaaaa-----taagggaagtalttggtctctacagaaaaaggaaaca	1390
Db	1944	CATCCGGTGTTGATGCACCATCTTCCAGCAACAATTTTCGGTGCTCCTCAGCGCTGGAACAG	1885
OY	1391	aggccccgaaaagtgaaagaccgggaatgagatctctttaaaanaacttggatgattactcng	1450
Db	1884	TTAAACCATTACGCCGACGTCAACTGCTTTTGTGATTACATCAAGAAGTGCGGCTAACAAATTATG	1825
OY	1451	tccagctatgacctgtttcttcgatctaaacagtgctcgaatgaaactatccaaacca--g	1507
Db	1824	TTCATTTCACAACCAATTGCAGACGCTCATTAAGAATACGATACGATGGAATAATGTAACCT	1765
OY	1508	aataattgggttatgaaacctgcacaactatgatatgtctcgaaggcaglatgctaacaaty	1567
Db	1764	ACAACCTGGGGTTATGATCACCAAACCTAATACGCCGCCGAACCACTGTTTTCAACTAAATC	1705
OY	1568	cgaabtgaaatgccc---gtataaaggatttaaggaaatgtcttcaaccctatcg	1624
Db	1704	CAGATGATCCAGCTCAGGTCAATTCGTGTGATTTGAAGATGATGGTTCAAGCTTATTCACGATG	1645
OY	1625	aaccacattgggtttaaacatgtagtgttgcataaacatacctttgccacogcaaatctcty	1684
Db	1644	CGGGTATTGGAGTCATTATGTGATGATGATGATGATTAACCAATACCTTCACAGTGTGATGAC	1585
OY	1685	acttcgataaaaaatgtaccacaatatatttacogtaagatgatcca-gtaattatacca	1743
Db	1584	CATTTCACAAACACAGTCCCTCATTTACTATTAATTCGATTAATTCACAGATGCTACCTTCACA	1525
OY	1744	accgatcacagtaacggaatgaaatltygaancungnaagagccaatlgttccaaabaattatt	1803
Db	1524	ATGGAAACGGGTGTGGAAATGAACACAGCAG-TGAACACGAATATTTTCGCAAGATATATG	1466
OY	1804	atgatactccctaagratltggttcaaagtatcatcatatttgacgcgtctccgttttgactta	1863
Db	1465	ATTGATTCCTCTTATATACGTGGGTGAGGAATATATATTGACGGCTCGGTTCGTTGACCTG	1406
OY	1864	atggcgcgtgccttgaaaaagacagatgtccaaagctgcctcggagctcaatgctattaat	1923
Db	1405	ATGGGGATTCATGTATGTCMAACACCATCAGATGATTCCTCAAAAGCTTGAGTAAAATTCGAC	1346



XX 06-MAY-1997: 97US-0045723.  
 PR (DUPO) DU PONT DE NEMOURS & CO E. I.  
 PA Broglie KE;  
 PI WPI: 1999-034728/03.  
 DR P-PSDB: W81973.  
 XX New nucleic fragments encoding corn pullulanase - useful to alter  
 PT pullulanase levels and thus starch structure in a host cell  
 PS Claim 2; Page 29-31: 42pp: English.  
 XX This sequence encodes a fragment of a novel *Zea mays* pullulanase protein  
 CC This protein and the encoding nucleic acid fragments may be used to  
 CC create transgenic plants in which corn pullulanase is present in cells at  
 CC levels differing from the norm, or in cell types or developmental stages  
 CC where it is not normally found, in order to alter starch structure in  
 CC those cells. They may also be used to isolate cDNAs and genes encoding  
 CC homologous pullulanase from the same or other plant species, to amplify  
 CC those genes and to produce antibodies for screening of cDNA expression  
 CC libraries.

SO Sequence 2904 BP: 867 A; 562 C; 674 G; 801 T; 0 other;

Query Match 2.7%; Score 76.2; DB 20; Length 2904;  
 Best Local Similarity 53.6%; Pred. No. 1.1e-11;

Matches 231; Conservative 0; Mismatches 185; Indels 15; Gaps 3;

OY 1503 ccaagaatattgggtatgacccctgcgaactatgattctcctgaaggcagatgctac 1562  
 DB 1033 ccctataatggtggtataaccctgtgtggtggtgcgtccaaagaagatgacgaag 1092  
 OY 1563 aatgcaatggtatgctgtataaagaatgtaagaatggtcttcacacccatcg 1622  
 DB 1093 taaccagatggtccaaagtcgtatcattgacggcgtatggtcagagccttgatcg 1152  
 OY 1623 tgaacaatggtgtatacgtatgtctataatacattacccttgcga----- 1672  
 DB 1153 cttaggtcttcaggtgtcgtatggtgtatatacattacccttgcgaagtgcccttc 1212  
 OY 1673 --cgcaaatctgactcgtataaattgtaccagaatattattaccgttcgatgaccca 1730  
 DB 1213 tgcctatcattcgtcttgcaagaatgtatcctgcgtatcaccctcagaag--gactct 1270  
 OY 1731 gtaattatccaacgagatcaggtactggaatgaaatgtcangcngaaagccaatgct 1790  
 DB 1271 aatggtcagactggaacagcgccgctgaaacaatacagaagtgagcatttc-attggt 1329  
 OY 1791 tcaaaaatttatattgattcccttaagtaattcggtaacagatgatatattgagcgctc 1850  
 DB 1330 tgaatagattaatcgltgactccttcgaattggcgagtaataacaaagttgacgggtc 1389  
 OY 1851 ccgtttgacttaatggtcgtctgtgaaaaaagacagatgtccaaagctccttcgagcgt 1910  
 DB 1390 cagattgacttaatggaatcagatcagaaaaagacaattgattagagcaaatcgctct 1449  
 OY 1911 tcaatgcatca 1921  
 DB 1450 tcaaaagcctta 1460

RESULT 14  
 X91106  
 ID X91106 standard: DNA; 5058 BP.  
 AC X91106;  
 XX 15-NOV-1999 (first entry)  
 DT  
 XX

DE Group B Streptococcus (GBS) antigens encoding DNA (clone 4).  
 XX  
 KW Group B Streptococcus; GBS; antigen; vaccine; Streptococcus infection;  
 KM sepsis; meningitis; pneumonia; immunocompromise; diabetes; liver disease;  
 KW cancer; veterinary; mastitis; ss.  
 XX  
 OS Streptococcus sp.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 1..666  
 FT /tag- a  
 FT /note- "see Y27354"  
 FT 763..1347  
 FT /tag- b  
 FT /note- "see Y27355"  
 FT 1362..1742  
 FT /tag- c  
 FT /note- "see Y27356"  
 FT 2266..5058  
 FT /tag- d  
 FT /note- "see Y27357"

W09942588-A2.

PD 26-AUG-1999.

PE 17-FEB-1999; 99MO-CA00114.

PR 20-FEB-1998; 98US-0075425.

PA (BIOC-) BIOCHEM VACCINS INC.

PI Boyer M, Brodeur BR, Charlebois I, Hamel J, Martin D;  
 RI Rioux C;

DR WPI: 1999-540309/45.

DR P-PSDB: Y27354, Y27355, Y27356, Y27357.

PT Novel group B Streptococcus antigens - useful as vaccine  
 PT compositions for prophylaxis or therapy of Streptococcus infections  
 PS Claim 10; Fig 4A; 154pp; English.

CC The invention provides Group B Streptococcus (GBS) antigens (Y27336-370)  
 CC and nucleic acids (X91103-X91111) encoding the antigens. The GBS antigens  
 CC can be recombinantly expressed using standard recombinant methodology.  
 CC The GBS antigens of the invention can be used as vaccine components for  
 CC the treatment or prophylaxis of diseases and symptoms mediated by  
 CC Streptococcus infection, especially group A Streptococcus (*S. pyogenes*),  
 CC GBS or *S. agalactiae*, *S. dysgalactiae*, *S. uberis*, *S. novae*, as well as  
 CC *Staphylococcus aureus*. The vaccines are administered to those individuals  
 CC at risk of GBS infection, particularly pregnant women and infants for  
 CC sepsis, meningitis, and pneumonia, as well as immunocompromised  
 CC individuals, such as those with diabetes, liver disease or cancer. The  
 CC vaccines also have veterinary applications, such as for the treatment of  
 CC mastitis in cattle. The present sequence represents a DNA encoding 4  
 CC GBS antigens of the invention.

SO Sequence 5058 BP; 1744 A; 816 C; 966 G; 1531 T; 1 other;

Query Match 2.7%; Score 74.4; DB 20; Length 5058;  
 Best Local Similarity 51.1%; Pred. No. 4.8e-11;

Matches 230; Conservative 0; Mismatches 208; Indels 12; Gaps 2;

OY 1508 ataattgggttatgacctgcgaactatgattctcgtgaaggcagatgcatcaaatg 1567  
 DB 4325 acatttgggtctatgacctcagacagagcatttcttcttggatgattcagagaac 4384  
 OY 1568 cgaatggt---aatgctcgtataaagaatgtaagaatggtcttcttcacatcgtg 1624  
 DB 4385 caaaagatcattcagccagcagatcgccgaattaaacaataatcattatcattacaac 4444

```

QY 1625 aacacatggggttaacatgatgttctctataatcaatcactcttccacgcgaatcctctg 1684
    ||| ||||| ||| ||||| ||| ||| ||| |||
Db 4445 gllggcatgggggttatacttgatgctgtctataatca-----cactgcacaaactt 4495

QY 1685 acttcgataaattgtacagaaatatttaccgtaacatgataccaggaatataacca 1744
    ||| ||| ||| ||| ||||| ||| ||| ||| ||| ||| |||
Db 4496 atctcttgagagatacagaacctaatatcattacattatgaataagctgtgtccacca 4555

QY 1745 cgaatcagactcgtgaatgaatgcangcngaaaggccaatggtttcaaaaatlatala 1804
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4556 gagaaagtttggagggggaacgtttaggaacacacatgcgaatgagtgctgtgtttgg 4615

QY 1805 ttgattccctaagaatggtggtcaatgagtcacataatgaagcttcgcttgcacttaa 1864
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4616 ttgattccataaataatccttaacaaagttgaattgattgttccgctttgatataga 4675

QY 1865 tggcctcgttggaaaagacaga-gtccaaagctgcctcggaagcttccatataatc 1924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4676 tgggagatcagatcgatcgctgcgattgaattagcttataaagaagctaaagctataatc 4735

QY 1925 caggaattgcacttaccgttgaagccatgga 1954
    || ||| ||| ||||| |||
Db 4736 ctatatagtattatgattggtgagggctgga 4765

RESULT 15
X33179
ID X33179 standard; cDNA; 784 BP.
XX
AC X33179;
XX
DT 06-JUL-1999 (first entry)
XX
DE Pullanase type partial length potato DBE cDNA clone SEQ ID NO:16.
XX
KW Potato; isoamylase debranching enzyme; DBE; starch; ss.
XX
OS Solanum tuberosum.
XX
PN WO9912950-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-GB02665.
XX
PR 06-SEP-1997; 97GB-0018863.
XX
PA (NMTT ) NAT STARCH & CHEM INVESTMENT HOLDING COR.
XX
PI Jobling SA, Schwall GP, Westcott RJ;
XX
DR MPI; 1999-229220/19.
XX
DR P-PSDB; Y04387.
XX
PT New isolated potato isoamylase-type debranching enzyme gene
XX
PS Example 2; Fig 10; 72pp; English.
XX
CC The present invention describes a novel nucleic acid sequence which is
CC obtainable from potato plants and carries at least a portion of an
CC isoamylase-type debranching enzyme (DBE) gene. Constructs containing the
CC nucleic acid sequence can be used to alter the starch properties of
CC plants such as potato, sweet potato, maize, wheat, barley, oat, cassava,
CC pea or rice. The starch can have increased branching and/or shorter
CC chain length, reduced peak viscosity, higher setback viscosity or
CC increased viscosity onset temperature. By using an antisense sequence
CC with greater homology to the native gene, greater inhibition can be
CC achieved. The present sequence represents a pullanase type partial
CC length potato DBE cDNA clone from the present invention.
XX
SQ Sequence 784 BP; 209 A; 161 C; 185 G; 229 T; 0 other;

```

```

Query Match 2.6%; Score 73.6; DB 20; Length 784;
Best Local Similarity 54.4%; Pred. No. 3e-11;
Matches 198; Conservative 0; Mismatches 151; Indels 15; Gaps 2;

QY 1508 ataattgggttatagccctcgcgaactatgatgttcttgaaggcagtatgtctacaatg 1567
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 ataattgggttatatactctgtctgtggtgaggttccaaagggaagctcgtcgtatag 472

QY 1568 cgaatgtaatcgtctgtatataaagaagttlaagaaatggttcttcactccatcgtgaac 1627
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 caaatgttcctgtcgtatcgtatcgttgaatttcggaatggtgcagggcaacttaatcgtatg 532

QY 1628 aacttggggttaacatggaatggttctctataatcactcttgc-----cagcg 1675
    || ||| ||||| ||| ||| ||| ||| ||| |||
Db 533 gtctacgtgtgtgattgattgttcttacaatcattacaatggaatggttccctttagtg 592

QY 1676 aaatctcgtactcgtataaattgttaccgaatatattacacgtacgaatgccagtaa 1735
    || ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 acttctcgttctagaagaagatcgttccaggttacttcttaaggagaatgctgactgagt 652

QY 1736 ttataccaacggaatcagtlacttgaatgaatgcaatgcaangcgaagccaatggtcaaa 1795
    || || ||| ||| ||| ||| ||| ||| ||| |||
Db 653 gcaattgaaa---atagcaaatgtgtataacaatccgcagtgagcatlctcatgttgagc 709

QY 1796 aattatattgattcccttaagtaattggtgcaatgagatcatatcatgacggttcggt 1855
    || ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 710 gcttaattcttgagaccctcaagtgctgagtcagttcatataaagattgaggtcttcggt 769

QY 1856 ttga 1859
    ||||
Db 770 ttga 773

```

Search completed: April 22, 2001, 06:56:18  
 Job time: 2716 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2001, 06:19:06 ; Search time 100.19 Seconds

(without alignments)  
4869.148 Million cell updates/sec

Title: US-09-262-126C-1

Perfect score: 2794  
Sequence: 1 gatgggaacacgacacgacat.....cgttaaaagtaataagaaaa 2794

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 302621 seqs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2786.4	99.7	4464	1 US-08-474-140-8	Sequence 8, Appli
2	2786.4	99.7	4464	1 US-08-474-140-9	Sequence 9, Appli
3	2786.4	99.7	4464	1 US-08-477-630-8	Sequence 8, Appli
4	2786.4	99.7	4464	1 US-08-477-630-9	Sequence 9, Appli
5	2786.4	99.7	4464	1 US-08-472-293-8	Sequence 8, Appli
6	2786.4	99.7	4464	1 US-08-472-293-9	Sequence 9, Appli
7	2786.4	99.7	4464	1 US-08-474-545-8	Sequence 8, Appli
8	2786.4	99.7	4464	1 US-08-474-545-9	Sequence 9, Appli
9	2786.4	99.7	4464	1 US-08-478-341-8	Sequence 8, Appli
10	2786.4	99.7	4464	1 US-08-478-341-9	Sequence 9, Appli
11	2786.4	99.7	4464	3 US-08-996-733-8	Sequence 8, Appli
12	2786.4	99.7	4464	3 US-08-996-733-9	Sequence 9, Appli
13	2786.4	99.4	2784	1 US-08-474-140-10	Sequence 10, Appli
14	2786.4	99.4	2784	1 US-08-477-630-10	Sequence 10, Appli
15	2786.4	99.4	2784	1 US-08-472-293-10	Sequence 10, Appli
16	2786.4	99.4	2784	1 US-08-474-545-10	Sequence 10, Appli
17	2786.4	99.4	2784	1 US-08-478-341-10	Sequence 10, Appli
18	2786.4	99.4	2784	3 US-08-996-733-10	Sequence 10, Appli
19	72.2	2.6	2646	1 US-08-365-189-7	Sequence 7, Appli
20	72.2	2.6	2982	1 US-08-365-189-2	Sequence 2, Appli
21	72.2	2.6	2988	1 US-08-365-189-1	Sequence 1, Appli
22	68.2	2.4	3437	4 US-08-860-339-17	Sequence 17, Appli
23	47.6	1.7	2244	1 US-08-476-519-10	Sequence 10, Appli
24	47.6	1.7	2244	5 PCT-US95-09323-10	Sequence 10, Appli
25	47.6	1.7	2334	1 US-08-476-519-1	Sequence 1, Appli
26	47.6	1.7	2334	5 PCT-US95-09323-1	Sequence 1, Appli
27	45.6	1.6	2712	2 US-08-410-784A-1	Sequence 1, Appli

C 28	44.2	1.6	7218	1 US-08-232-463-14	Sequence 14, Appli
29	43.4	1.6	2625	6 5457037-4	Patent No. 5457037
30	43.4	1.6	3336	6 5457037-1	Patent No. 5457037
31	41.8	1.5	2253	6 5457037-2	Patent No. 5457037
32	35.8	1.3	492	4 US-08-860-339-23	Sequence 23, Appli
33	34.6	1.2	1908	4 US-08-845-238-50	Sequence 50, Appli
34	34.6	1.2	1991	4 US-08-845-238-4	Sequence 4, Appli
35	34.6	1.2	2720	2 US-08-007-107-3	Sequence 3, Appli
C 36	34.4	1.2	2784	1 US-08-474-140-10	Sequence 10, Appli
C 37	34.4	1.2	2784	1 US-08-477-630-10	Sequence 10, Appli
C 38	34.4	1.2	2784	1 US-08-472-293-10	Sequence 10, Appli
C 39	34.4	1.2	2784	1 US-08-474-545-10	Sequence 10, Appli
C 40	34.4	1.2	2784	1 US-08-478-341-10	Sequence 10, Appli
C 41	34.4	1.2	2784	3 US-08-996-733-10	Sequence 10, Appli
C 42	34.4	1.2	4464	1 US-08-474-140-8	Sequence 8, Appli
C 43	34.4	1.2	4464	1 US-08-474-140-9	Sequence 9, Appli
C 44	34.4	1.2	4464	1 US-08-477-630-8	Sequence 8, Appli
C 45	34.4	1.2	4464	1 US-08-477-630-9	Sequence 9, Appli

#### ALIGNMENTS

RESULT 1  
US-08-474-140-8  
; Sequence 8, Application US/08474140  
; Patent No. 5721127  
; GENERAL INFORMATION:  
; APPLICANT: DEMER, PHILIPPE  
; APPLICANT: AMORY, ANTOINE  
; TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
; TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
; STREET: 2000 K Street, N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474.140  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: William F. Gagliano, Esq.  
; REFERENCE/DOCKET NUMBER: 37,136  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 429-0625  
; TELEFAX: (202) 429-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO. 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4464 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: genomic DNA  
; US-08-474-140-8  
Query Match 99.7%; Score 2786.4; DB 1; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 gatgggaacacgacacgacatgtccacataatttgcctcgtgtgattatcaacct 60

|||||  
Db 1250 GATGGCAACGACGACGATCATTTGTCACCTATTTTTGGCTGCTGGTGAATTCACACCT 1309  
QY 61 LGAGGTCATLGGALYGGCCAAAGACGAGGtggggctgaatacgaLlccaaLcaacccg 120  
Db 1310 TGGAGCTCATAGCATGTGGCCAAAGACGAGAGCTGGGGCTGAATTCGATTCATCAACCCG 1369  
QY 121 ggcgaaccccttggagctgttgcgaagtgctgataLlccaaagaaacccaagtcaagtga 180  
Db 1370 GCTGACCTCTTTGGAGGTGGTGAAGTGTGATATTCAGAGAAACCCAGTCAGGTGGA 1429  
QY 181 attacgtctgcgaactcaagaLltagaacaagaLgtgagcgttgacgctacatagattta 240  
Db 1430 ATTATCTCTTGCCACTCAAGATTGGACCAAGATGTGAGCCCTGACCCCTACATGATTTA 1489  
QY 241 agcaaaagaaalgaagtgctgctgtagaaagaaacagcaaatLlLlaaLgaaaaa 300  
Db 1490 AGCAAAAGAAATGAGTGTGGCTGTGAGAAAGAAACGCCAAATTTTAAATGAAAAA 1549  
QY 301 gatgctgaagatgagagctaaacccgctgtgaagcaagcttaLtagatgcttcaaccag 360  
Db 1550 GATGCTGAGATGCAAGCTAAACCCCTGTAAAGCAACCTTATTAAGTGTTCAAACCCAG 1609  
QY 361 gctgctgctaaacttagcagcgtctaaactctggggaagaaagcggcttaagct 420  
Db 1610 CTGCTGCTTAACTTACCACGCCCTTAACCTTGGGGAAGNNNAACGGCTTTACGGTT 1669  
QY 421 caLgaagcaacagcaaaLaaagalaLlccagtgacalcLglgaaggaLtgaaagtctgt 480  
Db 1670 CATGACACACACAGCAAAATTAAGATATTCCAGTCATCTGGAAGATGCAAGCTTGGT 1729  
QY 481 caagaLgtaaacgctgttctggagagtaactLccaataLlLltagagctccgaLtg 540  
Db 1730 CAAGATGTAAACCGTGTGTTGGAGGTACTCTCAACATATTTTGGAGGTCCGATTTG 1789  
QY 541 gcaactgataLtaacagtaactLltaaaaggtgactaaacatLlccatLlccatLlcca 600  
Db 1790 GCACCTGATATCACACTACTTTATTAAGAGGTGACTTAACAATCTCTATCAATTTCTCA 1849  
QY 601 ggaagatLlccctgaagaaacLtaacaaLlaaaagLtgctLlaaaLgaLlaagLtaaat 660  
Db 1850 GCGAGTCTTCCTGAAGCAAACTACCAATATTAAGTCCCTTAAATGATACCTGGAATTAAT 1909  
QY 661 cgaagttaaccatctgaacaatLltaaaagctccctgcggcggtgacaagcttaact 720  
Db 1910 CCGAGTTTACCATCTGCAACATTAATTAAGTCCCTCCGGGGTGGACAGGTCTACT 1969  
QY 721 tttctgataLlccgtccactcaLtgagctatgacaaatLlaaLtaLlccatLtgagat 780  
Db 1970 TTTTCTGATATTCCTGCTCCTCATGCTATGACAAATTAATTAATTCCTAAATGGCGAT 2029  
QY 781 LtaaaagLtaaaagcgggtLlaaaaggaLlccglgaagLtaactLlaaggaagaaLlcca 840  
Db 2030 TTTCAAGTAAAGAGCGGGTTAAAGCGATCTCCTGACGGTACTACTAGCGGAAGATCCA 2089  
QY 841 gatgtagagcaLtaactctgtccatLltaaaagatLggtcaLtaagcaaaagcgggtgatacc 900  
Db 2090 GATGTGAGCCATCTGCTCATTTCAAGAGATGGCTATAGGCAAGCGATGATACCT 2149  
QY 901 cgtaaLgtgcttaactcaLtaacagtaactatLtaaggaatgatactctggaaataactat 960  
Db 2150 CGTAATCTGCTTAATTCATACAGTACTACTATTTCAGAGATATGCTTTGGGAATACCTAT 2209  
QY 961 acaacagaagcaaaacLltaaaagLlggcaacaaactLltaactcaagLlaaLgtLct 1020  
Db 2210 ACACAGAAACCAACACCTTTAAAGTCTGGGACCAACCTTCTACTCAAGTAATGTTCTT 2269  
QY 1021 ccttaLgaagcgtgcaagggctctgtcaaaaaaactgctactatgacgcaLcggggcat 1080  
Db 2270 CTTTATGACAGTGCACAGGGGTCTGTAAACAAAATCGTACTATGAGGATGGGGCAT 2329  
QY 1081 ggtgtctggaagcaagcgtLtaaaacaaactLgaaaatLgfatLtaactatgata 1140  
|||||

Db 2330 GCTGTGCGAAGCAACGGTTAATCAAAAACCTTGAAATTTGGTATTCANGTATGAGTA 2389  
QY 1141 acaagccaagcctLtaaccgaacggctLgtatccctLtaLgaactLgcatLgcaaccaat 1200  
Db 2390 ACAGGCCAAGGCTTACCAGAGGGCTGTGATCTTATGCAACTGGGATTTGCACCAAT 2449  
QY 1201 ggaaggaagcaLgtatLgtgacctLgtctaaacaaactcctgtcgtcgtgaagatgat 1260  
Db 2450 GGAAGGAGAGCATGATTTGTGACCTGGCTTAAACATATCTGCTGGCTGGACAGTAT 2509  
QY 1261 aaacatLlccgccaagaagaaLataagaatLgagttcLlaLgaaaLgagatLlccgagac 1320  
Db 2510 AAACATATTTACGCAAAAGAAATATGAAGATGAGGTCAATTAAGAAATGATGTCCGTGAC 2569  
QY 1321 LtlccatLgaacccLtaactcgggtatgaaaataaaaggaatLlLtgacctLtaagaa 1380  
Db 2570 TTTTCCATTTGACCTTAATTTGGGTATGAAAAATTAAGGAAGTATTTGGCTTACAGAA 2629  
QY 1381 aaagaaacaaagggccctLgaacagLaaagaaggggaLagatLlccLtaaaacaaactggg 1440  
Db 2630 AAAGAAACAAAGGGCCCTGACAACTAAAGCGGGGATGATTTCTTAAACAACTTGGG 2689  
QY 1441 attacatLgtLtaagctLtagcctLgtLlccgcatLtaacagLgtcgaLgaacatLgaca 1500  
Db 2690 ATTACTATGTTCAAGCTTATGCTGTGTTTGCATCTAAACGTGTGATGAATGATGATCCA 2749  
QY 1501 acccaagaLtaatLtgaggLltaLgaacccLgcacaaLtaLgtLlccLgaagggcaLagct 1560  
Db 2750 ACCCAAGATATTTGGGTATGACCTCGCAACATATGATGTCTGAAGGGCAATATGCT 2809  
QY 1561 acaaaLtgcaatLgtatgtctgtataaaagatLlaagaaatLgtLlccLtaaccat 1620  
Db 2810 ACAAATCGAAATGTAATGCTCGTATAAAGATTAAAGAAATGATTTCTTCACTCAT 2869  
QY 1621 cgtLgaacaaLlLgggtLtaaaLgaLgtLgtLtaaaLtaactLlLgcgaacaaat 1680  
Db 2870 CGTAAACATTTGGGCTTAATGACCTCGCAACATATGATGTCTGAAGGGCAATATGCT 2929  
QY 1681 LctgactLcgaLtaaaatLgtLtaaccagaatLataLtaacgLaagatLtaaccagLtaata 1740  
Db 2930 TCTGACTTCGATTAATTTGTCAGAAATATTAATACCGTACAGATGACAGTAAATATA 2989  
QY 1741 ccaacgaLtaagtaactgtgaatLgaaaLlLtgangcngaaagggcaatLgtLlcaaaatL 1800  
Db 2990 CCAAGGATCAGATGACGGAATTAATTTGCANCCNAAAGGCCAAATGTTCAAAAATTT 3049  
QY 1801 attatLlccactLtaagLlLlLggLtaLgaatLtaataLlLtagagctLlccgctLlLlgac 1860  
Db 3050 ATTATTAATTTCCCTTAAGTATTTGGGTCAATGAGTATATTAATTTGACGGCTTCCGTTTGAC 3109  
QY 1861 LtaatLgcgtLgtLlLgaaaagacgaLglLccaaagctLgcLcgaagctLcaLgtatL 1920  
Db 3110 TTAATGCGCTGCTTGGAAAAAGACAGATGTCNAACCTGCTCGAGCTTCATGCTAAT 3169  
QY 1921 aatLcaggaatLtgacLlLlccgtLgagcaLtgagcgtLggaacLlLlLgcactLgcagat 1980  
Db 3170 AATCCAGAAATTCGACTTTACGGTGAGCCATGAGCGGGTGAACCTTGCACTCCAGAT 3229  
QY 1981 gatcaagctLlLgaacaaaggaagctLcaaaaggaLlLgggaLlLaggggtLlLtaagaat 2040  
Db 3230 GATCAGCTTCTGACAAAAGAGCTCAAAAAGGCAATGGAGATGCGGTAAAGACAAT 3289  
QY 2041 LtaagaagcgtLtgagcggcaatLgtLlLgtatLlLlccgctLcaagLlLlLtgagcaag 2100  
Db 3290 TTTAGAAACGGTTGAGAGGCAATGTCTTTGATTTCCCTCAAGTTTGGACACAGT 3349  
QY 2101 gcaacaggtLtaactLgaatLgaatLgaatLgagcgtLgaagggagatLtaatLgaactLacc 2160  
Db 3350 GCAACAGGCTTAACCTGATGAATTAAGAAATGAGCTTGAAGGAGATTAATTAAGACTTACC 3409  
QY 2161 LtlLtaacaggtLgaagaatLtaactatLgtcaaaagLcaLgaLtaactLcacccLlLtgagac 2220  
Db 3410 TCTTCACAGGTGAGCAATTAATTAATGATGTCACAAAGTATGATTAATCAACCTTTGGGAC 3469  
|||||

QY	2221	aaataagccttaagcaatccctaagtctccgaagcgatcgattaaatgataaactc	2280
Db	3470	AAATATGACCTTAAGCATCTCTAATGATTTCCGAAGCGATTCGATTAAATGATGAATC	3529
QY	2281	gcacaagcagltgtltaagactacaaagcglttccatcatbtaagcagcg99gaagaaatg	2340
Db	3530	GCACAAACGATGTTATGACCTCACAAGCGTTCCATTATGCAAGCGGGGAAGAAATG	3589
QY	2341	cttcgtanaaaagcg9gcaagcaaaatgataatbtaagcagcgatcggtcaatgagttt	2400
Db	3590	CTTCGTANAAAAAGCGGCAACGCAATAGTTTAAATGACGCGATGCGTCATGAGTTT	3649
QY	2401	gattggagcag9gaagctcaatctcaagtlgtttcaactatataaggggtactaac	2460
Db	3650	GATTGGAGCGAGAAAGCTCAATATCCAGATGTTTCAACTATTATTAAGCGGGCTAAATCCAC	3709
QY	2461	cttcgcttctatcacccagccttcgcgataagcaagcaatcgtaaatcaatagccactc	2520
Db	3710	CTTCGCTTATATCACCCAGCCTTCGCCATGACGACAGCTAATGAATCATATGCACTC	3769
QY	2521	caattccataaatagtcacagagaaacagatg9ccatbtaataactgtaactgtaataaa	2580
Db	3770	CAATCTCTAATATAGTCCAGAGAACACAGAGCCCTATGATTAATCATGATCATGTTAAATAA	3829
QY	2581	gacaaatagg9gaatatcatgttgttttaaacccaataaanaactgtaagcaacctaat	2640
Db	3830	GACAAATAGGGGAATATCATGTTGTTTATTAACCCAAATAAATACGTGACCAACCATTCAT	3889
QY	2641	ttgacgagcg9gaatatlg9gcaatcaatgtctacagagcg9gtaagtaggaatccacctt	2700
Db	3890	TTGCGAGCGGGGAATATGCGCAATATGCTACGAGCGGTAAAGTAGGAATCCACCTT	3949
QY	2701	ggttcaagcagag9gaatgtgtccaagtacaggtatatactatgataccttcatcaagag	2760
Db	3950	GGTCAACCAAGGAAGTGTCCAAAGTACCAAGGTATATATGATGATCCTTCATCAAGAG	4009
QY	2761	gtaagccagcaagccaggtataaagtatatgaaaa	2794
Db	4010	GTAAGCCCAAGCACCGGTAAAAAGTAAATGAGAAA	4043

RESULT 2  
 US-08-474-140-9  
 Sequence 9, Application US/08474140  
 Patent No. 5721127  
 GENERAL INFORMATION:  
 APPLICANT: DEWEER, PHILIPPE  
 APPLICANT: AMORY, ANTOINE  
 TITLE OF INVENTION: PULLINAWASE, MICROORGANISMS WHICH  
 TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREP  
 TITLE OF INVENTION: PULLINAWASE AND THE USES THEREOF  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 STREET: 2000 K Street, N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/474,140  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wilhem F. Gadiano, Esq.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-41

```

? TELECOMMUNICATION INFORMATION
? TELEPHONE: (202) 429-0625
? TELEFAX: (202) 293-1850
?
? TELEX: 650 383-5605
?
? INFORMATION FOR SEQ ID NO.: 9:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4464 base pairs
? TYPE: nucleic acid
? STANDEDNESS: single
? TOPOLOGY: linear
?
? MOLECULE TYPE: genomic DNA
?
? OS=08-474-140-9

```

Query Match	99.7%;	Score 2786.4;	DB 1;	Length 4464;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2793;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	gatgggaacagcaacaagatcattccaacltttccgcgcggtgattatcaact	60
Db	1250	gatgggaacagcaacaagatcattccaacltttccgcgcggtgattatcaact	1309
QY	61	tggagctatgtagtggccaaagaacggaggtgggctgtaacagattccaacccg	120
Db	1310	tggagctatgtagtggccaaagaacggaggtgggctgtaacagattccaacccg	1369
QY	121	gctgaccttttggagcgtttgcaagtgctgtatccaggaaccgaagtcagtagga	180
Db	1370	gctgaccttttggagcgtttgcaagtgctgtatccaggaaccgaagtcagtagga	1429
QY	181	attatcgtctgcgaacccaagatttgacccaagaatgtgagcgtcgacccgtacataagatla	240
Db	1430	attatcgtctgcgaacccaagatttgacccaagaatgtgagcgtcgacccgtacataagatla	1489
QY	241	agcaaaaggaaatagagtgctgtgtagaaggaaacagccaaatttttaacagaaaa	300
Db	1490	agcaaaaggaaatagagtgctgtgtagaaggaaacagccaaatttttaacagaaaa	1549
QY	301	gatgtctggagatgagcgtcaaaacccgcttlaagcaacgctattatgagtcttcaaaccg	360
Db	1550	gatgtctggagatgagcgtcaaaacccgcttlaagcaacgctattatgagtcttcaaaccg	1609
QY	361	gtctgtggttlaaactttagccagccgttlaactctctgggggaagmnnaagcggcttlaacglt	420
Db	1610	gtctgtggttlaaactttagccagccgttlaactctctgggggaagmnnaagcggcttlaacglt	1669
QY	421	catgacgcgaacacgcaaaataaggaatattccagtgacaactctgtgaagatgcaagtcctgtgt	480
Db	1670	catgacgcgaacacgcaaaataaggaatattccagtgacaactctgtgaagatgcaagtcctgtgt	1729
QY	481	caagaatgtaaacccgctgttttggcaagtgacccctccaaacatatttggaggtccgatttg	540
Db	1730	caagaatgtaaacccgctgttttggcaagtgacccctccaaacatatttggaggtccgatttg	1789
QY	541	gcaacctgtaataacagtaacttattctlaaaaaaaggtygaactaaacatctatcaatltcta	600
Db	1790	gcaacctgtaataacagtaacttattctlaaaaaaaggtygaactaaacatctatcaatltcta	1849
QY	601	ggagagctcttctctgaaaggaactacccaataaagtgtgctttaaatagatagcctggaatac	660
Db	1850	ggagagctcttctctgaaaggaactacccaataaagtgtgctttaaatagatagcctggaatac	1909
QY	661	cggagtttaacccactcgtgaacaacataattaaagaaccccgcgcggtggtgcacaacttcat	720
Db	1910	cggagtttaacccactcgtgaacaacataattaaagaaccccgcgcggtggtgcacaacttcat	1969
QY	721	ttttcgtatattcgtlccactcatgtgaacttctatgacacaacatlaatatcttaatcgagat	780
Db	1970	tttttctgtatattcgtlccactcatgtgaacttctatgacacaacatlaatatcttaatcgagat	2029
QY	781	ttacaagtacaaagcggggtttaaaacgagatctcgttgaagtgtaactcttaggggaagatcca	840
Db	2030	ttacaagtacaaagcggggtttaaaacgagatctcgttgaagtgtaactcttaggggaagatcca	2089

QY	841	ggtctgagacatactctccatctcaacaagagtgatcatcaggcaaaagagtgatacc	900
Db	2090	gatttgagacatactctgttcattcaaaagatgcttatcagccaaagcggatgatacct	2149
QY	901	cgtaatgcttaattcatcacaagtaactatctcagaagatgactctggaaataactat	960
Db	2150	ccttaattgcttaattcatcacaagtaactatctcagaagatgactctggaaataactat	2209
QY	961	acacaggaagcacaaccccttaaaagctcgggcaacaactctactcaagttaattgtctct	1020
Db	2210	acacaggaagcacaaccccttaaaagctcgggcaacaactctactcaagttaattgtctct	2269
QY	1021	ccttaatgacgtgcaacccggtctcgttaacaacaactctactcaagttaattgtctct	1080
Db	2270	ccttaatgacgtgcaacccggtctcgttaacaacaactctactcaagttaattgtctct	2329
QY	1081	ggctgtgtgggaagcaacccggtctcgttaacaacaactctactcaagttaattgtctct	1140
Db	2330	ggctgtgtgggaagcaacccggtctcgttaacaacaactctactcaagttaattgtctct	2389
QY	1141	agaagcccaagagctctcaacccggtctcgttaacaacaactctactcaagttaattgtctct	1200
Db	2390	agaagcccaagagctctcaacccggtctcgttaacaacaactctactcaagttaattgtctct	2449
QY	1201	ggaagcagagagatgattgtggaacccggtctcgttaacaacaactctactcaagttaattgtctct	1260
Db	2450	ggaagcagagagatgattgtggaacccggtctcgttaacaacaactctactcaagttaattgtctct	2509
QY	1261	aaacatactcaagcccaaaagatatagaagatgaggtcactatgaatgtgtcgtgac	1320
Db	2510	aaacatactcaagcccaaaagatatagaagatgaggtcactatgaatgtgtcgtgac	2569
QY	1321	tttccatctgaccccttaattcgggtatgaaaaataaagggaatatttgctctctaaaga	1380
Db	2570	tttccatctgaccccttaattcgggtatgaaaaataaagggaatatttgctctctaaaga	2629
QY	1381	aaaggaaacaaaggccctctgacacgttaaaagcgggagatgattcctctaaacaactctg	1440
Db	2630	aaaggaaacaaaggccctctgacacgttaaaagcgggagatgattcctctaaacaactctg	2689
QY	1441	attactcatgtctcagctctctgtctctctgacatacaagctgtcgaatgaactgtca	1500
Db	2690	attactcatgtctcagctctctgtctctctgacatacaagctgtcgaatgaactgtca	2749
QY	1501	accceaaataatctggggtctgacacccctgcaactatgactgtcctcgaaggcagatgct	1560
Db	2750	accceaaataatctggggtctgacacccctgcaactatgactgtcctcgaaggcagatgct	2809
QY	1561	acaaatgacagatgataatgctcgtatbaaaagatttaaggaaatggtctctcaactcat	1620
Db	2810	acaaatgacagatgataatgctcgtatbaaaagatttaaggaaatggtctctcaactcat	2869
QY	1621	gctgaagacatactggggtctgacatgataatgctcgtatbaaaagatttaaggaaatggtctctcaactcat	1680
Db	2870	gctgaagacatactggggtctgacatgataatgctcgtatbaaaagatttaaggaaatggtctctcaactcat	2929
QY	1681	ctcgaactcgaataaactgtacacagaatacttaacccgaagatgacagatgaatactata	1740
Db	2930	ctcgaactcgaataaactgtacacagaatacttaacccgaagatgacagatgaatactata	2989
QY	1741	ccaacgagatcaggtactcgtgaatgaaatgcaatgcaagcgaaggccaaatgtctcaaaaatt	1800
Db	2990	ccaacgagatcaggtactcgtgaatgaaatgcaatgcaagcgaaggccaaatgtctcaaaaatt	3049
QY	1801	attatgtatcgccttaagatgaggtgcaatgaggtatcatatgagcaggtctcgtttgac	1860
Db	3050	attatgtatcgccttaagatgaggtgcaatgaggtatcatatgagcaggtctcgtttgac	3109
QY	1861	ttaaatgagctgctctgaaaaagacagatgctcaaaagctgctcgaagctcatgactat	1920
Db	3110	ttaaatgagctgctctgaaaaagacagatgctcaaaagctgctcgaagctcatgactat	3169

ADDRESS: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/477,630  
APPLICATION NUMBER: US/08/477,630  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-42  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SRO ID NO: 8  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-477-630-8

Query Match 99.7% Score 2786.4; DB 1; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gatgggaacgacgaacgacatcttccactatcttgcctcgtggtgattatcaacct 60  
DB 1250 GATGGGAACGACGAACGATCATGTCACATATTTTGGCGTGTGATATCAACCT 1309  
QY 61 tggagctatgagatgtggccaaaagagagtggtggcgatgaatcaatcaacg 120  
DB 1310 TGGAGCTATGAGATGTGGCCAAAAGAGGAGGTGGGTGAATACATTTCAATCAACG 1369  
QY 121 gctgacctcttggagctgttgcaagtgctgatatccaggaacccaagtcagtaga 180  
DB 1370 GCTGACCTTTTGGAGCTTTTGCAGTCTGCTGATTTCCAGGAACCCAGTCAAGTAGA 1429  
QY 181 attatcgctgcacactcaagaattggaccaaagatgtgagcgtgacccgtacatagatta 240  
DB 1430 ATTATCGTTCGCACATCAATGATGACCAAAAGATGTGAGCGCTGACGCTACATAGATTAA 1489  
QY 241 aacaaagaaatgagtggtgtgtgtagaagaagaacagccaattttatataatgaanaa 300  
DB 1490 AACAAAGAAATGAGTGTGCTGTGAGAGGAACAGCCAAATTTTATATATGAAAAA 1549  
QY 301 gatgtcgaagatgacagtaaacccgctgtaagcaacgctatattagatgctcaaacag 360  
DB 1550 GATGCTGAGAGATGACACTAAACCCGCTGTAGAGCAAGCTTATTATGCTTCAAAACAG 1609  
QY 361 gtgctggttaaaccttagcagccgcttaactcttggggaagmnnaagcggtcttcaggtt 420  
DB 1610 GTGCTGCTTAACTTAGCCAGCGCTTAACTCTTGGGAAGGNNAAGCGGCTTTACGGTT 1669  
QY 421 catgacgacacgaacaaatagatatccagtagacatctgtgaagaatgcaagcttgat 480  
DB 1670 CATGACGACACGACAAATAGATATTCAGTGAACATCTGTGAAGATGCAAACTCTTGGT 1729  
QY 481 caagatgtaaccgctgttctggcaggtaccttccacaatalatttggaggttcagattg 540  
DB 1730 CAAGATGTAACCGCTGTCTTTGGCAGGTACCTTCCAAACATATTTTGGAGGTTCCGATTGG 1789

QY 541 gacactgataatcacagtaacttattaaaaaagggtactaacatctcatcaatctca 600  
DB 1790 GCACCTGATTAATCACAGTACTTATTAAAAAGGTGACTAACATCTCATCAATTCTCA 1849  
QY 601 ggaagatcttccctggaaggaaacacacataataagtgctttaatgatatgctggaataat 660  
DB 1850 GGAGATCTTCCCTGAAAGAAACATCAATATTAAGGTGCTTTAAATGATAGCTGGAATTAAT 1909  
QY 661 ccgaattaccatcttgacaacatlaattlaacagctcccgccggcggtgacacgctact 720  
DB 1910 CCGAGTACCCTCTGACACATTAATTATACAGTCCCTCGCGGCGGTGACACGCTCACT 1969  
QY 721 ttcttgatattccgcgtccatcatgagctctatgacacaaatlaattactaatgctgata 780  
DB 1970 TTTTCTGATATTCCTGCTCCATCATGATGACACATTAATTAATCTTAATGCGGAT 2029  
QY 781 ttacaagtagaaagcggggtttaaaacggatctcgtgaaggttaacttgaagggaagatcca 840  
DB 2030 TTACAAAGTAGAAAGCGGGGTAAACGGATCTCTGACCGTTTACCTTAGCGGAATCCA 2089  
QY 841 gatgtgacatactctgtccattcaaacagatggtctatcaaggcaagcaggtgataact 900  
DB 2090 GATGTAGCCATTACTCTGTCCATTTCAACAGATGGCTATCAGGCAAGCAGTGTATACCT 2149  
QY 901 cgtaatgtcttaattcatcaacagtaactatctcagagaatgacttgggaatacctat 960  
DB 2150 CGTAATGTGCTTAATTCATACACAGTACTACTATTCAGGAGATGATCTTGGAAATACCTAT 2209  
QY 961 aacagaaagcaacaaactttaagctctgggacccaacttctactaaatgtaatgtctct 1020  
DB 2210 ACACAGAAAGCAACACCTTTAAAGTCTGGGACCAACTCTTACCTCAAGTAAATGTCTT 2269  
QY 1021 cttaatgacagtgacaggggtctgtgaacaaaatcgttaactatgacggcatctgggcat 1080  
DB 2270 CTTTATGACAGTGCACACGGGTTCTGTACAAAATGCTACTATGACGGCATGGGCCAT 2329  
QY 1081 ggtgtgtgggaagcaacgggttaatacaaaccttgaanaattgttatatagtatgagta 1140  
DB 2330 GGTGTGTGGGAAGCAACGGGTATTAATCAAAACCTTGAAATGTGATATACATGTATGAGTA 2389  
QY 1141 acagggcaagggcttaccgggaagggctgtgtgaccttaatgcaactgcatgtaacaaat 1200  
DB 2390 ACAGGCCAAAGGCTTACCCGCAACGGCTGTGATCTTATGCAACCTCGATTCACCAAAAT 2449  
QY 1201 ggaacgagagggcatgattgtggaacctggcctaaacagatctcgtgcttggaaacagtgat 1260  
DB 2450 GGAACGAGAGGCTATGATTGTGGACCTGGCTAAACAGATCTCTGCTGGTGAACATGAT 2509  
QY 1261 aaacataactaagcacaagaatagagaatgaggtcaactatgaatggaatgctgcgtgac 1320  
DB 2510 AAACATATTAACGCCAAAGAAATATAGAAATGAGGTCAATCATGAAATGATGTCCTGAC 2569  
QY 1321 ttcttcattgacaccttaattccgggtatgaaaaataaagggaagatatttgctcttaagaa 1380  
DB 2570 TTTTCAATTGACCTTAATTCGGGTATGAAAAATTAAGGGAATGATTTGGCTCTTACAGAA 2629  
QY 1381 aaagaaacaaagggccctgacaaagctlaaagacggggaatagatctcctlaaaacaaacttgg 1440  
DB 2630 AAAGAAACAAAGGGCCCTGACAAAGTAAAGAGGGGATGATTCCTTAAACAAACTTGGG 2689  
QY 1441 attactcaatgctacgtatgctgttctgcgaactcaacagtgctgataaactgtacca 1500  
DB 2690 ATTACTCATGTTCAAGTTATAGCTGTGCTTGCATCTTACAGTGTGATGAAATCTGATCCA 2749  
QY 1501 acccaagaataatgagggtatgaacctcgaaactatgatttctctaaaggcgagtgtgct 1560  
DB 2750 ACCCAAGATTAATTTGGGTTATGACCTCGCAACTATGATGTTCTCTGAAGGCGAGTATGCT 2809  
QY 1561 acaaatgcaaatgtaatgctgtatlaaagaagttaaaagaaatggttctcttaaccat 1620  
DB 2810 ACAAAATGCGAATGTATGCTGTATTAAGAAATGATTTAAGGAAATGTTCTTCACTGCAT 2869  
QY 1621 cgtgaacacatgtgggttaacatgagatgttctctataatcaactcttgcacagcaaatc 1680



Db 1490 AGCAAAAGAAATGAGGTGTGCTGTGTAGAGAGAAACACCCAAATTTTATATATATAAAA 1549  
Qy gatctgagatgagtcagcctaaaccgcgtgtaagcaagcttaattgatgcttcaaccag 360  
Db 1550 GATGCTGAGATGCGACGTAAACCCGCTGTAAAGCAAGCTTATTGTAGCTTCAAAACAG 1609  
Qy gtctggtgtaaaactlaagccagcgtttaaactcttgaggaaagannaagcggcttlaagct 420  
Db GTGCTGTTAAAGTTAGCCAGCCGCTTAACTCTTGGGAAAGNNNAAGCCGCTTACGGTT 1669  
Qy 421 catgacgacaaagcaataaagaatattccagtgacatctgtgaaagatgacaaagcttggt 480  
Db CATAGCGACAGCAAAATAGATATTCAGTACATCTGTGAAGATGCAAGCTTTGCT 1729  
Qy 481 caagatgtaaacgcgtgtttggaagtgacctccaacatattttggaggttcgactgg 540  
Db 1730 CAAGTGTAAACCGCTGTTTGGCAGGTACTTCCAACTATTTTGGAGGTTCCCATTTGG 1789  
Qy 541 gcaacctgataatcaagtaacttattaaagaagtgactaaacalctlatcaattctca 600  
Db 1790 GCACCTGATATTCACAGTACTTATTTAAAAAGGTGACTAAACAATCTCTATCAATTTCA 1849  
Qy 601 ggaagatcttctgaaagaaactaccatataaagaatgagctttaaagtaagctggaataat 660  
Db 1850 GGAGATCTTCTGGAAGAACTACCAATATTAAGTGGCTTTAAATGAATAGCTGGAATAT 1909  
Qy 661 ccgagttacccatctgacaacatatttaaacagtcacctgcggcggtgacacgtacct 720  
Db 1910 CCGAGTTACCATCTGACAACTTAATTTAACAGTCCCTGCCGGGGGTGACACGTCACT 1969  
Qy 721 ttctcgataltccgtccactcalgacgltclatgacacaatataatccatlaactgacggt 780  
Db 1970 TTTTCGTATATTTCCGTCACATCATGACGTATGACACAAATTAATTCCTAATGCGGAT 2029  
Qy 781 ttacaagtaagaagcggggtttaaagaagctcgtgacggttaactctgagggaagatcca 840  
Db 2030 TTACAGTAGAAGAAAGGGGGTTAAACGGATCTCGTACGGTTACTCTAGGGGAATCA 2089  
Qy 841 gatgtgaacatactctgtccatcaacaagatggtctacgaagcaagcaggtgtaact 900  
Db 2090 GATGTGACCATCTCTCTTCATTAAACAGATGCGTATACAGCAAGCAGGTGATACCT 2149  
Qy 901 cgtlaatglttaaltcalcaagtaactlaactlaactlaactlaactlaactlaactlaact 960  
Db 2150 CGTATGTGCTTAATTCACAGTACTATCTACAGGATGATCTTGGCAATACCTAT 2209  
Qy 961 acaagaaagcaaacactttaaagctcggggaacaaacttcaactcaagtaaatgtctt 1020  
Db 2210 ACACGAAAGCAACACTTTAAAGTGTGGCACCACTTCTACAGTAATGTCTT 2269  
Qy 1021 cttaatgacgtgcaagcgtgtctgtaacaanaatcgttaactatgacgcatcgagccat 1080  
Db 2270 CTTTATGACAGTGCAGACGGGTTCTGTAAACAAATCGTACATATACGCGCATCGGCCAT 2329  
Qy 1081 ggtgtgtggaagcaagcgttlaactaaaacacttgaanaatltggtlaltacatgtaagyla 1140  
Db 2330 GGTGTGTGGAAGCAAGCTTAACTCAAAACCTTGAAATTTGGTATTTACATGATAGGTA 2389  
Qy 1141 acaagccaaagcttcaaccgaaagcgtltgatacttcaactgacatgacccaat 1200  
Db 2390 ACAGGCCAAAGCTTACCCGAAACGGCTGTGATCTTATGCAACGCGATTCACCAAAAT 2449  
Qy 1201 ggaagcgaagcagatgtgtggaaccttgctaaacagatcctgctgagcgtggaagcgtgt 1260  
Db 2450 GGAAGCAAGCATATTTGTGACCTGGCTTAAACAGATCCCTGCTGGAGCAAGTGTAT 2509  
Qy 1261 aaacatatacgcacaaagataaagaatgagtcacatctatgaaatgtaagltccgtgac 1320  
Db 2510 AAACATATTACGCCAAAGCAATATAGAAGATGAGTCAATCTATGAATGATGTCCGTGAC 2569  
Qy 1321 ttctccatgacccctaatctcggtatgaaaaaaataaagggaagtaattggctcttaacaga 1380

Db 2570 TTTTCCATTGACCCCTAATTCGGGTATGAAAATTAAGGCAAGTATTTGGCTTACAGAA 2629  
Qy 1381 aaaggacaagaagccctgacaacgttaaagcagggagataagattcccttaaacacttgg 1440  
Db 2630 AAAGGACAAAGGCGCTGACACAGCTAAAGACGGGATAGATTCTTAAACAACTTGGG 2689  
Qy 1441 altactcalgttcaagcttatgctgttctcgacatcaagtgatgaaatgatacca 1500  
Db 2690 ATTACTCATGTTCACTTATGCGCTGTTTCCGATCTAACAGATGTGATGAACACTGATCA 2749  
Qy 1501 acccaagataatttgggtgtatgacccctcgcaactatgattgtccctgaaggcagatgct 1560  
Db 2750 ACCCAAGATAATTGGGGTTATGACCTCGCAACTATGATGTTCTGAAAGGAGATATGCT 2809  
Qy 1561 acaaatgcgaatggtgaatgctcgtataaagaagtttaagaaatgtgtcttccatccat 1620  
Db 2810 ACAATGCGAATGTGTAATGCTGTATAAAGAGTTTAAGAAATGCTTTCACCTCCAT 2869  
Qy 1621 cgtgacaacatlygggtlaacalgtgltgtlctatlaactlaactlttgccacgcaatlc 1680  
Db 2870 CGTGAACACATTTGGGGTTAACATGATGTTGTCTATTAATCATACCTTGGCCACGAAATC 2929  
Qy 1681 tctgacttcgataaanaattgtaccagaataatttaaccgtaagatgataccagttatata 1740  
Db 2930 TCTGACTTTCGATAAATTTGTACAGATATTTATACCGTACATGATGACGGTAAATTATA 2989  
Qy 1741 ccaacggaatcaggtactgtaaaatgaaatgcangcngaaagccaaatgtctcaaaaatt 1800  
Db 2990 CCAACGATACAGTACTGCAAAATGAAATGACANGCAAGGCAATGTTTCAAAAATTT 3049  
Qy 1801 altattgaltccctlaaglatatlyggltcaatgaglatcalatgacggtcttcglttgac 1860  
Db 3050 ATTATTTGATTTCCCTTAAGTATTTGGGTCAATGATGATCATATTTAGACGGCTTCCGTTTTCAC 3109  
Qy 1861 ttaatgagcgtctctggaagaacagatgtccaaagcgtccttcggagcttcatgtatct 1920  
Db 3110 TTAATGGCGCTCTTGGAAAGAACAGATGTCAAAGCTCGCTCGAGGTTTCAATGCTATT 3169  
Qy 1921 aatccaagaaatgcacttlaagcgtgagccaatggaacggtggaacacctcagctgccaagt 1980  
Db 3170 AATCCAGAAATTTGCACTTTACGTTAGCCATGACCATGACCGGTGGAACCTCTCCTACTGCCAGAT 3229  
Qy 1981 galcagctctgacaagaagagctcaaaaaggcaltgaggatgacggtgttlaatgacaal 2040  
Db 3230 GATCAGCTTCTGACAAAAGACGCTCAAAAAGCATGAGTGAAGGGGTGTTAATGACAAAT 3289  
Qy 2041 ttaagaaacggttggaaagcaatgtcttgaattcttccgtccaaggttttcgacaagt 2100  
Db 3290 TTACGAAACGCGTTGGACGGCAATGTCTTGAATCTTCCGCTCAAGGTTTCCGACAGGT 3349  
Qy 2101 gcaacagcgttaactgataatgaatgaatgagcgttgaaggagatlaaiaacttacc 2160  
Db 3350 GCAACAGGCTTAACGTATCATTAAGATGCGCTTGAAGGAGATTTAATCACTTTACC 3409  
Qy 2161 tcttaccaggtgagacaatlaactatgcaagaatgatalaactacaccccttggagc 2220  
Db 3410 TCTTCACACGAGTGAAGCAATTAATATGTCACAAAGTCATGATTAACACCTTTGGAC 3469  
Qy 2221 aaaaatagccctaaagcaatcttaatgattccgaagcggatcggatlaaataatgatacc 2280  
Db 3470 AAAAATAGCCCTTAAGCAATTCCTAATGATTTCCGAAGCGATCGGAATTAATGATGAATCT 3529  
Qy 2281 gcaacagcaggtgttatgacctcaagaagcgttccatlaatgcaagggggaagaatg 2340  
Db 3530 GCACAAAGCAGTTGTTATGACCTTCACAAAGCGTTTCCATTTATGCAAGCGGGAAGAAATG 3589  
Qy 2341 ctctglanaaaagcggcaacgacaatlaatlaatgcaaggcagatgacatgacatgac 2400  
Db 3590 CTTTCGTANAAAAGCGGCAACGACATATTAATGAAGGAGGATGCGGTCAATGATTT 3649  
Qy 2401 gatgtgagcaggaagcctcaatataccagaatgtttcaactatataaggcgttaacac 2460  
Db 3650 GATTGGAGAGGAAGCTCAATATTCAGATGTTTTCACCTATTATAGCGGCTAATCCAC 3709



QY 2461 ctcgctcgtatcaccagcccttcgcgatacgaacgaactaataatcaatagaccctc 2520  
|||||  
Db 3710 CTTCGTCGTGATCACCGACGCTTCGCCATGACGACACTAATGAATCAATAGCACCTC 3769  
QY 2521 caattcctaataatgctcagaagaacacagtgctctatgaatlaactgcatcgttaataa 2580  
3770 CAATCTCTAATATGTCCAGAACACAGTGCCTATGAATTAATCTGATTAATAA 3829  
QY 2581 gacaaatgggaaatacatctgttcttatacaaccaataaactagacaacacaaat 2640  
|||||  
Db 3830 GACAAATGGGAAATATCATTTGTTTATACCCAAATAAATCTAGCAGACCATCAAT 3889  
QY 2641 ttcgagcgagggaataggcaatcaatgctacgagcggtaaggtaaggaaatccaccctc 2700  
3890 TTGCCGACGCGGAAATGCGCAATCAATGCTACGACGCGTAAGGTAGAGAAATCCACCTT 3949  
QY 2701 gctcagcagaagggaagtgctcgaagtcaccggtatatactctgtagtccctcacaagag 2760  
3950 GGTCAACGACAGAGGAACTGTCCAAAGTACCAAGTATATCTATGATGATCTTCATCAAGAG 4009  
QY 2761 gtaagccagacacacggtataaagaataagaaaa 2794  
4010 GTAAGCCGACACACGCTAATAAGTATAGAAA 4043  
Db

RESULT 5  
US-08-472-293-8  
Sequence 8, Application US/08472293  
Patent No. 5731174  
GENERAL INFORMATION:  
APPLICANT: DENEER, PHILIPPE  
APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOBER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,293  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William F. Gagliano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-44  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ. ID NO.: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: Linear  
MOLECULE TYPE: genomic DNA  
US-08-472-293-8

Query Match 99.7%; Score 2786.4; DB 1; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;

	Matches 2793;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 1	gatggagacacgaacacgacatctgtccactatcttggccctcgtgtgattcaacct 60				
Db 1250	GATGGAGACACGACACGATCATTTGTCACATATTTTGGCCCTGCTGTGATTATCAACT 1309				
QY 61	tggagctatgagtgtggccaagaacgaggtggggctgtaatacgaatttcaatcaacg 120				
Db 1310	TGGAGCTATGATGATGTGGCCAAAAGACGAGGTGGCGCTGAATACGATTTCAATCAACCG 1369				
QY 121	gctgacctcttggagctgtgtcgaagtgctgatatccagaagaacccaagtcaagtgaga 180				
Db 1370	GCTGACCTCTTTGGAGCTGTTGGCAAGTGTGATATTCCAGGAAACCCAACTCAGTAGGA 1429				
QY 181	attacgtctgcacatcaagatltgaccaagaatgtgagcgtcgaacgctacatagattta 240				
Db 1430	ATTATCGTTTGGCACTCAAGTTTGACCAAAAGTGTGAGCCCTGACCCCTACATAGATTTA 1489				
QY 241	agcaaaaggaaatgaaggtgtggtctgtgagaagaacacagccaatlttlaataatgaaaaa 300				
Db 1490	AGCAAAAGAAATGAGCTGTGCTGTGAGAAAGAAACAGCCAAATTTTATTAATGAAAAA 1549				
QY 301	gatgtgagatgacgtataaaccgctgtgaagaacagcttatatgagtgcttcaaacag 360				
Db 1550	GATGCTGAGATGCAAGCTTAACCCGCTGTAGCAACGCTTATTTAGATGCTTCAAACAG 1609				
QY 361	gtgctgttaaaccttaacgacgttcaactcttgaggagagmnaacgagcttaacgct 420				
Db 1610	GTGCTGTTAACCTTAACCGCGCTTAACCTTGTGGGAAAGNNNAACGCGCTTTAGCGTT 1669				
QY 421	catgaacacacgaacaaataagatatccagtgacalcgtgaaagatgcaagltcgtgt 480				
Db 1670	CATGACGACACAGCAAAATAGATATCCAGTGCACTGTGAAGATGCAAGCTTGTGT 1729				
QY 481	caagatgtaacccgtgtgttggagaggtacccccaataatcttggagctcagattg 540				
Db 1730	CAGATGTAAACCGTGTGTTGGAGGTGACCTTCCAAATATTTTGGAGGTCCGATTTG 1789				
QY 541	gcacctgataatcacagtaacttatataaagaagtgcatacaactcctcaacttca 600				
Db 1790	GCACTGATATATCACACTACTTTATTAATAAAGGTGATCAATCTATTAATGCGAT 1849				
QY 601	ggagatcttccgtgaagaaactaacatataaagtggtctttaaagtacgtggaataat 660				
Db 1850	GGAGATCTCTCTGAAGGAACTACCAATATAAGTGCCTTTAAATGATACCTGGAATAAT 1909				
QY 661	ccgaattaccatcttgacaacatttaataagagtcctcccgaggggtgacaacgttact 720				
Db 1910	CCGAGTTACCCATCTGCAACATTAATTTANACGTCCCTCCGCGGTGCAACGCTACT 1969				
QY 721	tttcgtatactcgttcactcatgacgtctatgacacaacttaataatcccaatgcygat 780				
Db 1970	TTTTTCGATATTTCCGTCTCACTCATGCGTCTATGACCAATTAATATCTAATGCGAT 2029				
QY 781	ttacaagtagaaagcgaggtttaaagcagatcgtgtgaggttlaactlaagggaaatcca 840				
Db 2030	TTACAACTAGAAAGCGGGGTTAAACGAGATCTGTGACGCTTACTCTAGGGGAAATGCA 2089				
QY 841	gatgtgagccatactcgttcacttcaagaagatgggcatcaaggcaaaagcggtgatccct 900				
Db 2090	GATGTGAGCCATACTCTGTCATTCAACAGATGGCTATAGGCAAAAGCAGGTGATACCT 2149				
QY 901	cgtaaatgtcttaattatcatcagttactatctcaggaatgtagctctggaaataacct 960				
Db 2150	CGTAATGTCTTAATTCATACAGACTACTATTTCAGAGATGATCTTTGGGAATACCTAT 2209				
QY 961	acacagaagaacaaacacctttaaagcttggcaccacacacctctactcaagttaagtctc 1020				
Db 2210	ACACAGAAACCAACACCTTTAAAGTCTGGCACCAACTTCTCAAGTAATGTCTT 2269				
QY 1021	cttataagagtgcaacggtgtctgtacaaaaaatcttactatgaacggtacgtggacat 1080				
Db 2270	CTTTATGACAGTGCACMGGGTCTGTATACAAAAATCTGATGAGGAGCATGGGCCAT 2329				



QY 1081 ggtgtgtggaagcaagcgtttaaataaaccttgaaatgtgtatataatgtagtga 1140  
|||||  
Db 2330 GGTGTGTGGAAGCAAGCAGTGAATCAAAACCTTGAATAATGGTATGATGAGGTA 2389  
QY 1141 acagcgcaagagcttaccggaagcggtgtgtatgcttataatgtaactgtgacccaat 1200  
|||||  
Db 2390 ACAGCGCAAGAGCTTACCAGCAAGCGGTGTGTGATCTTATGCAACTGCCATTTGCACCAAT 2449  
QY 1201 ggaacgaagagcatgaltgtgacgttgcaaacagatccctgtgcttggaacagtat 1260  
|||||  
Db 2450 GGAACGAGAGCATGATGTGTGACCTGGTAAACAGATCCTGGTGGTGGAAACAGTAT 2509  
QY 1261 aaacatactacgccaagaataatagaagatgaggtcatctatgaatgtatgtcgtgac 1320  
|||||  
Db 2510 AAACATTTACGCCAAGAAATATAGAAATGAGTCAATCTTGAATAATGATGCTCGTAC 2569  
QY 1321 ttctccatgaccttaattctgggtatgataaataaagaagagtagtattggccttacaaga 1380  
|||||  
Db 2570 TTTTCCATTCACCTTAATTCGGGTATGAAAAATAAAGGGAAGTATTTTCGCTTACAGAA 2629  
QY 1381 aaagaaacaaagggcccttgacaacgtlaaagaaggagtagatctcttaaacacatggg 1440  
|||||  
Db 2630 AAAGAAACAAAGGGCCCTGCACAAAGTAAAGCGGGGATGATTTCTTAAACAACTTGG 2689  
QY 1441 attactaatgtcaggtctatgctgttttcgcatctaacagtgctgatgaaactgtatcca 1500  
|||||  
Db 2690 ATTACTATGTTGACGTATGCTGTTTTCGATCTAACAGTGTGATGAAGAACTGATCCA 2749  
QY 1501 acccaagataattgtgggttatagaccctcgcaactatgtagtctctgaaaggcagtatgct 1560  
|||||  
Db 2750 ACCCAATATATTTGGGTATGAGCCCTCGCAATGATGATGTTCTGAAAGGGCAGATGCT 2809  
QY 1561 acaatgcgaatgtgtaatgtcgtlataaagaagttlaagaaatgtgtcttctacatcat 1620  
|||||  
Db 2810 ACAATGCGAATGTGATGCTCGTATAAAGAGTTTAAAGAAATGCTTCTTCACTCAT 2869  
QY 1621 cgtgaacacatgtgggttaacatgtagtgtgctataatacatccttgcacgcaaatc 1680  
|||||  
Db 2870 CGTGAACACATTTGGGGTTAATCATGATGATGTTGTATATCATCTTACCTTGCACGCAATC 2929  
QY 1681 tcttacttgataaattgttcacgaatatataatccgtlaagatgatacgaatata 1740  
|||||  
Db 2930 TCTTACTTCGATTAATAATTTGACAGAAATTTATTAACCTGATGATGACAGTAAATTA 2989  
QY 1741 ccaacgagatcaggtactggaatgaatltgcangcngaagccaatgtgtccaaatlt 1800  
|||||  
Db 2990 CCAACGGATCGATGATGGAATGAATTTGCANCGNAAAGCCAAATGCTTCAAAAATTT 3049  
QY 1801 attacttacccttaagtagtgggtcaatgtagtataatltgacggtctcgtttgac 1860  
|||||  
Db 3050 ATTATTTGATTCCTTAAGTATTTGGGTCAATGATGATCATATTTGACGGCTTCCGTTTAC 3109  
QY 1861 ttaatgctgctgttgaagaaagacagatgfcgaagctgcttcggaactcgaatgtat 1920  
|||||  
Db 3110 TTAATGGCGCTGCTTGAAGAAACACAGATGTCCAAAAGCTGCTCGAGCTTCATCTATTT 3169  
QY 1921 aatccaggaattgcaacttactgtagcagatgagcaggttgaaactcgcacatgcaat 1980  
|||||  
Db 3170 AATCCAGGAATTTGCACTTACGTCAGGCATGAGCGGTGAACCTCGCACTGCAAT 3229  
QY 1881 gatcagcttctgacaaaaagagctccaaaagcattggagtagcggtgtttaaagaat 2040  
|||||  
Db 3330 GATAGCTTCTGACAAAAGAGGCTCAAAAAGCATGGAGTAGGGGTGTTTAATGACAAAT 3289  
QY 2041 ttacgaagacggttggagcaggaatgtcttcttctccgtaaggtttggagcaagt 2100  
|||||  
Db 3290 TTACGAAGCGGTTTGAGCGGAATGTCTTGTGATCTTCCGCTCAAGGTTTGGCGACAGT 3349  
QY 2101 gcaacaggtcttaactgtagcaatlaagaatgagcgttgagggaglatlaagaacttacc 2160  
|||||  
Db 3350 GCAACAGGCTTAACGTGATGCAATTAAGAAATGGCGTTGAGGGGAGATTAATGACTTTAC 3409

QY 2161 tcttcaccaggttgagacaattaactatgtcacaagtgcatgataactacaccccttggac 2220  
|||||  
Db 3410 TCTTCACCAGGTGAGACAAATTAATGATGTCAAAAGTCAATGATTAATACACCTTTGGAC 3469  
QY 2221 aaataagcccttaagcaatcccttaatgattccgaaggggtcgtgattaaatgtagaactc 2280  
|||||  
Db 3470 AAAATAGCCCTTAAGCAATCTTAATGATTTCCGAAGCGGATCGGATTAATAATGATGAATC 3529  
QY 2281 gcaacagagttgtlatagcttcaacaagcgcttccatctcatgacaagcggggaagaatg 2340  
|||||  
Db 3530 GCACAGCAGATGTTTATACCTCAACAGCGCTTCATTCATGACAAAGCGGGGAAGAAATG 3589  
QY 2341 ctctgtanaaaagcgcgcaagcaaatagttataatgcaaggcgtatgcygtaatgattc 2400  
|||||  
Db 3590 CTCTGTANAAAAAGCGGCAAGACAAATGATTATATGACAGCGGATGCGTCAATGATTT 3649  
QY 2401 gattggagcaagaaagcccaatccaaatgttttcaacttlaagcgggttaatccac 2460  
|||||  
Db 3650 GATTGGAGCAGGAAAGCTCAATATTCAGATGTTTCAACTATTAATGAGCGGCTTAATCCAC 3709  
QY 2461 ctctgcttgaacccaagccttccgcatgacagcagcctaataatgaatagccactc 2520  
|||||  
Db 3710 CTCTGCTTGAATCAACCGACCTTCGCGATGACGACGATTAATGAATATGACCACTC 3769  
QY 2521 caattccctaaatagtcagagaaacagtggtcctatgtaataactgcatgattataaa 2580  
|||||  
Db 3770 CAATTCCTAAATAGCCAGAAACACAGTGGCTGTGATTAATCAATGATCATGTTAATTA 3829  
QY 2581 gacaatgtgggaatatcatgtgttattataacccaataaactgtgcaacacatcat 2640  
|||||  
Db 3830 GACAAATGGGGAATATATCTTGTTTATTAACCAATAAATCAATGACCAACCTTCAT 3889  
QY 2641 ttgcccagcggaatgtggcaatcaatgctcagcagcgtlaaggttagaagatccaccc 2700  
|||||  
Db 3890 TTGCGGACGGGAATGGGCAATCAATGATCTAGAGCGGTAAGTGAAGATCCACCTT 3949  
QY 2701 gttcaagcagaggaagtgtccaaagtlaccaggtatataatgtagtacttccatcaag 2760  
|||||  
Db 3950 GGTCAAGCAGAGGGAAGTGTCAAGTACAGGATATCTATGATGATGATCTTCATCAAG 4009  
QY 2761 gtaagccgaagccaaggtlaaagaatagaana 2794  
|||||  
Db 4010 GTAAGCCAGACCAAGGTAAAGTAATGAGAAA 4043

RESULT 6  
US-08-472-293-9  
Sequence 9, Application US/08472293  
Patent No. 5731174  
GENERAL INFORMATION:  
APPLICANT: DEMER, PHILIPPE  
APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,293  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Wilhelm F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-472-293-9

Query Match 99.7%; Score 2786.4; DB 1; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggagacacgaacacatcatttccactattttgcccctgctggtggtatcaacct 60  
DB 1250 GATGGGAAACACGACACATCATTTTGGCCCTGCTGCTATATCAACCT 1309  
QY 61 tggagctatgagctggtgccaagacgaggtggtggtggtatcaacctcaacg 120  
DB 1310 TGGAGCTATAGAGTGTGGCCAAAGAGGAGGTGGGCTGATATCAATCAACCG 1369  
QY 121 gctgacaccttggagctgttgcgaagtgctgatatccagaagaaacccaatcggttga 180  
DB 1370 GCTGACCTTTTGGAGCTGTTGCAAGTGTGATATTCAGGAAACCCAGTCAGGTAGGA 1429  
QY 181 atatcgttcgcactcaagaatlgaaacaaagatlgagcgtlgaacgctacatagattta 240  
DB 1430 ATTATCCTTGCGACTCAAGATTTGACCAAGATGTGAGCCCTGACCTCATATGATTTA 1489  
QY 241 agcaaaagaaatgaggtgtgtgtgtgtgaaagaaacagcaaatcttctaataagaaaa 300  
DB 1490 AGCAAAAGAAATGAGGTGTGTGTGAAAGAAACGCCAAATTTTATATGAAAAA 1549  
QY 301 gatgctgaagatgcagatgaacccgctgtgaagcaacgcttatatgaatgcttcaaacag 360  
DB 1550 GATGCTGAGATGCAGCTAAACCCGCTGTAAAGCAACCTTATTTAGTCTTCAAAACAG 1609  
QY 361 gctcgtgttaaaacttaacgacgctltaactcttggaggaagaaagcggcttaccgct 420  
DB 1610 GTCCTGCTTAACTTACCAGCCGTTAACTCTTGGGAGGNNNAACGGCTTTACCGTT 1669  
QY 421 catgagacacgaacaaataagagataatccagtgacatctgttgaagatgcgaatcttgc 480  
DB 1670 CATGACACACAGCAAAATTAAGATATTCAGTGCATCTGTGAAGATGAGAGCTTGGT 1729  
QY 481 caaagatgaacgctgttctggcaggtgacttgcgaacataatcttggaggcttccgattg 540  
DB 1730 CAAGATTAACCGCTGTTTGGCAGGTACTCTTCCACATATTTTGGAGGTTCCGATTGG 1789  
QY 541 gcaacgtataatcagatctatctataaaagggtgaactaacaatctctatcaatctca 600  
DB 1790 GCACCTGATATCACTACTTATTAATAAAGGTGACTTAACAATCTCTATCAATTTCTCA 1849  
QY 601 gggagatcttccgtaagaaatcacaataaagtggtcttaaatgaatcgtgaataat 660  
DB 1850 GGGAGATCTTCTGAAGAAATCAACATATTAAGTGGCTTAAATGATACCTGGAAATAT 1909  
QY 661 ccgaatccacatctgcaacatctaatttaacagttccctgcgagcgttgaacgcttacc 720  
DB 1910 CCAGATTACCATCTGCAACATTAATTTAAAGTCCCTCCGGCGGTGCACAGCTACT 1969  
QY 721 ttctcgtataatccgtccactcagctcagctatagacacaaatlaataatccatctgagat 780  
DB 1970 TTTTCGTAATTCGCTGCACATCGAGTCTATGACACAAATTAATCTCTAATGGCGAT 2029

QY 781 ttaacagtagaaacggggttaaaacgagatcgtgagcgttactctcagggaagatcca 840  
DB 2030 TTAAAGTAGAAAGCGGGGTTAAACGAGATCTGTGACGGTACTCTAGGGGAATCCCA 2089  
QY 841 gatgtgagccatctcgtccatctcaaacagatggtcatcagaacaaagcagggtataact 900  
DB 2090 GATGTGACCATCTCTGTCATTAACAAGATGGCTATCAGGCAAAAGCGATGATACCT 2149  
QY 901 cgtlaatgcttaatcactacagtaactatctcaagagatgatatcttggaaataactat 960  
DB 2150 CGTAACTGCTTAATTCATACAGTACTATTCAGAGATATATCTTGGGAATACCTAT 2209  
QY 961 acacagaagcaaacacactttaagtcctggcaccacacttctaactcaagtaaatgtctct 1020  
DB 2210 ACACAGAAACCAACACCTTTAAAGTCTGGGACCAACCTTCACTCAAGTAATGTTCTT 2269  
QY 1021 cttatagacagtgcaacgggttctgttaacaaaaatcgttaactatgaacgcatcggccat 1080  
DB 2270 CTTTATGACAGTGCACAGGGGTCTGTAAACAAAATGCTATGACCGCATCGGCCAT 2329  
QY 1081 gctgtgtggaagcaacggttcaatcaaaaccttgaaaatlgatlatcatgataagata 1140  
DB 2330 GGTGTGGGAGACCAACGCTTAATCAAAACCTTGAAATTTGATTTATCATGATGAGTA 2389  
QY 1141 acagccaaaggtctcaccggaacggtgtgtatccttgaacatcgatlgacacaaat 1200  
DB 2390 ACAGGCCAAAGCTCTACCCCAACGCGCTGTGATCTTATCAACTGCGATGACCAAAAT 2449  
QY 1201 ggaagagaagagcatgatatgtgagaccttggttaaaacagatctctgctggaacagtgat 1260  
DB 2450 GGAACGAGAGGCTGATTTGTGACCTGGCTAAACACATCTCTGCTGGAACAGTAT 2509  
QY 1261 aaacataatcgcacaaagaaatagaaagatgaagtcacatcctgaatgataatgatacgtgac 1320  
DB 2510 AAACATATTCGCCCAAAAGATATAGAAAGATGAGTCACTATGAAATGATGTCGGTAC 2569  
QY 1321 ttctcattgacccctaatctcgtgtgatagaataaaggaagatctgtccttcaagaa 1380  
DB 2570 TTTTCCATTTACCCCTAATTTGGGTATGMAAAATTAAGGGAATTTTGGCTTACAGAA 2629  
QY 1381 aagaagacaaagggcccttgaacagctgaagagcggggaatgaatcccttaaaacactggg 1440  
DB 2630 AAAGGAACAAAGGCCCTTGCAACAGTAAACACGGGATACATTCCTTAAACAACTTGGG 2689  
QY 1441 attactatgttcaagcttcaagctgttctgcacatcaaacagtgatcgaatgaactgtacca 1500  
DB 2690 ATTACTATGTTCAGCTTATGCTGTTCGATCTTAAAGTGCATGATGATGATGCA 2749  
QY 1501 acccaagataatctgggtgtatgaacctcgaactatgagatgttccctgaagggcagatgct 1560  
DB 2750 ACCCAAGATTAATTTGGGTTATGACCCCTCGCACTATGATGTTCTCTGMAAGGCGATATGCT 2809  
QY 1561 acaaatgcaagatgaaatgtctgataaaagagtttaagaaatgtgtcttcaatccat 1620  
DB 2810 ACAAAATGCAATGTATGCTCTGTATTAAGAAATGTGTTCTTACATCCAT 2869  
QY 1621 cgtgaacacatctggggttcaacatgagatgtgtctatataatcaactcttgcacgcaaatc 1680  
DB 2870 CGTGAACACATTTGGGTTTACATGATGATGTTGTATTAATCATACCTTGGCACGCAATC 2929  
QY 1681 tctgactctgataaataatgtaacgaatatattacggtatcgatgataatcattata 1740  
DB 2930 TCTGACTTGCATTAATAATTTGACGAATATTAATTAACGTCATGATGAGTGAAGTAATTA 2989  
QY 1741 ccaacgcatcaggtactggaatgaaatgacangcnaaagacaaatggttcaaaaaattc 1800  
DB 2990 CCAACGATACGTAAGTGAATTAATTTCAACGCAACGCAATGTTCAAAAAATTT 3049  
QY 1801 ataatgattccctlaaglaatctgggtcaatgaatlatataatlgacggctcctgattgac 1860  
DB 3050 ATTATTAATTCCTTTAAGTATTTGGGTCAATGATGATCATATTTGACGGCTTCCGTTTGGAC 3109  
QY 1861 ttaatggcgtgcttggaaaaagacagatgccaagcgtcctcgaagcttcaatgataat 1920



|||||  
Db 1730 CAGAGTAAACCGCTGTTGGAGGTAACCTCCACATATTTTGGAGGTCGGATTGG 1789  
OY 541 gcaactgataatcaagtaacttataataaaaggtagactacaactcctatacacttca 600  
Db 1790 GCACCTGATAATCAAGTACTTTATTAATAAAGGTGACTAACAACTCTTCAATTCCTCA 1849  
OY 601 ggaagatcttcctgaaggaacatccaaataaagtgctttaaataagctgaataat 660  
Db 1850 GGAGATCTTCTTCAAGCAAACTACCAATATAAGTGGCTTTAAATGATTAACCTGCAATAAT 1909  
OY 661 ccaagttacacatctgacacatbaatttaacagttcccttcgagggatgacacgtact 720  
Db 1910 CCGAGTTACCACTTGACAAACATTAAATTAAACGTCCCTCCGGGGTGCACACGTACT 1969  
OY 721 ttctcgataatccgtccactcctatgactatgacacaaataaatacctaactgagat 780  
Db 1970 TTTTGTATNTCCGTCCACTCATGAGTATGACAACTTAATATCTTAATGGGAT 2029  
OY 781 ttaacagtaagaagcggggtltaaaagcctcgtgagggatlaactaagggaagatcca 840  
Db 2030 TTACAACTAAGAAAGCGGGCTTAAACCGATCTCTGACGGTACTCTAGCGGAAGATCCA 2089  
OY 841 gattgagcaatactctgtccatctcaaacagatggctatcaaggcaagcaggatgatacct 900  
Db 2090 GATGTAGCCATACTCTGTCTATTCAACAGATGCTATACAGCAAGCAAGGTGATACCT 2149  
OY 901 cgtatagtgcttaataatcatcacagtaactatactcagagaatgactctgggaatacct 960  
Db 2150 CGTAACTGCTTAATTCATCAGACTACTACTATTCAGAGATGATCTTGGGAATACCTAT 2209  
OY 961 aacacagaagaacacacacactltaaaagcttgagcccaactctactcaagttaagtctc 1020  
Db 2210 ACACAGAAACCAACACCTTTAAAGCTGCGGCACCAACTTCTACTCAAGTAAAGCTTCTT 2269  
OY 1021 ctctatagcagtgcaagggctctgtatacaaaaaatcgtaactatgacggatcagggcat 1080  
Db 2270 CTTTATACAGTGCACAGGGTCTCTGTAAACAAAATCGTACCTATGACGGCATCGGCCAT 2339  
OY 1081 ggtgtgtggaagcaacggtltaatacaaaccttgaataatgtgatacatgtatgagga 1140  
Db 2330 GGTGTGGGAGACAGCGTTAATCAAAACCTGAAATTTGTTATCATGATGATGAGSTA 2389  
OY 1141 acaaggccaaggtcttaeccgaagggctgtgtatccttaagcaactgcatlgaaccaat 1200  
Db 2390 ACAGCCCAAGCTCTACCCACAGCGCTGTGATCCCTTATCACTCCGATGCAACCAAT 2449  
OY 1201 ggaacgagagagcatgattgtgagaccttgactaaacagatcctgtgagctggaacagt 1260  
Db 2450 GGAAAGAGAGCATGATTGTGGACCTGGCTTAACAGATCTCTGCTGGCTGGACAGTGAAT 2509  
OY 1261 aacacatactacgccaagaatataagaagatgaggtlcaactatgaataatgagtgatc 1320  
Db 2510 AAACATATTACGCCAAAGATATATGAAGATGAGGTCTATGAAATGATGATGTCGGTAC 2569  
OY 1321 ttctccatlgaaacctaaatttgagggtatgaataataaagggaagatcttggtcttaacgaa 1380  
Db 2570 TTTTTCATTAACCTTAATTTGGGTATGAATAAATAAGGAAAGTAATTTGGCTTTACGAA 2639  
OY 1381 aaaggaacaaaggcccttgacaacgtaaaagcggggaataagatccttaaaacaacttgg 1440  
Db 2630 AAAGGAACAAAGGGCCCTGACAAAGTAAAGCGGGATTAATTCCTTAATAACAACTTGGG 2689  
OY 1441 attactcaatgctcaagcttatgctgttcttgcatcaacagtgctgaatgaactgata 1500  
Db 2690 ATTACTATGTTCAGCTTATGAGCTGTTTGCACTCTAACAGTGCATGAATGATGATCCA 2749  
OY 1501 acccaagaataatggggtatlgacccctcgacaactatgattctctgaaggcaglatgct 1560  
Db 2750 ACCCAAGATATATGGGTATGAGCCCTGCAACATATGATGTCTCTGAAGGCGCAATATGCT 2809  
OY 1561 acaaatcgaaatgtaagctcgtatataaagagtttaagaaatggtctctcaactcat 1620  
|||||

---

Db 2810 ACAATCGAATGCTAATGCTCGTATAAAGACTTTAAGNAATGCTTTTCACTCCAT 2869  
OY 1621 cgtgaacacacttggtgtataacatgatactgtctataatacactctgcacacacatc 1680  
Db 2870 CGTGAACACATTTGGGTTTAACATGATGTTGTATTAATCATACCTTTGCCACGCAATTC 2929  
OY 1681 tctgactctgataaattgttaccagaataltatlaacgtatcgatgataccagtaattata 1740  
Db 2930 TCTGACTTCATTAATAATTTGACCGAATATTATTAACCTACGATGATGACGATTAATATA 2989  
OY 1741 ccaaggaatcaggtactgggaatgaaatlgcangcnaaaagcccaatggttcaaaaatt 1800  
Db 2990 CCAAGGATACGACTACTGAAATGAATTTGCANCCNAAGGCGCAATGTTCAAAAATTT 3049  
OY 1801 attatgattcccttaagatattggttcaatgagatcataatlgacggtccgttctggac 1860  
Db 3050 ATTATGATTTCCCTTAAGTATTTGGTCAATGAGATCATATTAGAGGGCTTCCGTTTTGAC 3109  
OY 1861 ttaatggcgtctcttggaagaaagacagatgccaagctgctclogagctatgatact 1920  
Db 3110 TTAAATGCGGCTGCTTGGAAAAAGACACATGTCCAAACCTCCCTCGAGCTTCAATGCTATT 3169  
OY 1921 aatccaggaattgacacttctacgtgtagcactgagcgggtggaacctctgacatgcagat 1980  
Db 3170 AATCCAGGAATTCGACTTTACGGTAGCCATGACGGGTGACACTTGCACATGCAGAT 3229  
OY 1981 gatcagctcttgacaaagaagagctcaaaaggcactggaatgaggtgttataatgacaat 2040  
Db 3230 GATCAGCTTCTGACAAAGAGAGCTCAAAAAGCATGAGAGTGGAGTGGTTAATGACAAAT 3289  
OY 2041 ttacgaaacggttgagcggcaatgcttctgatactctccgtcgaaggttttggaacagct 2100  
Db 3290 TTACGAAGACCGGTGAGCGCAATGTCTTGAATCTTCCCTCAAGCTTTTGCAACAGCT 3349  
OY 2101 gcaacaggtcttaactgtatgcaatlaagatggcgttgagggagatataatgatacttacc 2160  
Db 3350 GCACACAGCTTAAGTATGCAATTAAGATAGCCTTAAAGGGAGATTAATGACTTTACC 3409  
OY 2161 tcttcaacaggtgagaacaataactaactatgacaaatgataatcaactacaccccttgag 2220  
Db 3410 TCTTACACAGGTGAGCAATTAATTAATGTCACAGTCAATGATTAACCTTACCCCTTTGGAC 3469  
OY 2221 aaaaatagccttaagcaatccttaatgattccgaagcggatcggatlaaatgatactc 2280  
Db 3470 AAAATACCCCTTAACCAATCTTAATGATTCGAAACCGATCGATTAATAAGATGAATC 3529  
OY 2281 gcaacagcagttgttatagacctcaagggcttccatctatgcaagcggggaagaattg 2340  
Db 3530 GCACACAGATGTTGATGACCTCAAGAGGCTTTCATTCATGCAAGCGGGGAAGAAATG 3589  
OY 2341 ctctgtanaaaaggcggcgaacgaatagttataatgcaagcgaatgctgtcactaagttc 2400  
Db 3590 CTTCTGTANAAAAGCGCGCAACGACAAATATTAATGATGATGATGATGATGATGATGAT 3649  
OY 2401 gattggagcgggaagctcaataatcaatgatttcaactatctaaacgggtctaaatccac 2460  
Db 3650 GATTGGAGCGAAGAGCTCAATATCCAGATGTTTCAACTATTAATTAACGGGCTTAATCCAC 3709  
OY 2461 ctctgtctatcaacccagccttcgcatgacgaacagctaaatgaatcaatgaacaccc 2520  
Db 3710 CTTGCTTATGATCAACGAGCTTCCGATGACGACACTAATGAATGAATGAATGAATGAAT 3769  
OY 2521 caatctcaatagtcagaggaacagatggcctaataatcaatgataatgataatgataat 2580  
Db 3770 CAATTCCTTAATATCTCCAGAGAACACAGTGGCTTAATTAATGATGATGATGATGATGAT 3829  
OY 2581 gacaaatggggaatatcatctgtgtataaaccaaataaataatgataaaccataat 2640  
Db 3830 GACAAATGGGAAATATCATTTGTTATTAACCAATTAATAATGATGATGATGATGATGAT 3889  
OY 2641 ttgcgaagcgggaatgagcaatcaatgataagcggatgaagtaggagaaatccaccc 2700  
Db 3890 TTGCGGAGCGGGAATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3949

OY 2701 ggtcaagcagaggaagtgcccaagaccagatatactatgatccttcacaaag 2760  
|||||  
Db 3950 ggtcaagcagaggaagtgcccaagaccagatatactatgatccttcacaaag 4009  
OY 2761 gtaagccagacacagtgtaaaagtaataagaaa 2794  
|||||  
Db 4010 gtaagccagacacagtgtaaaagtaataagaaa 4043

## RESULT 8

US-08-474-545-9

Sequence 9, Application US/08474545

Patent No. 5736375

GENERAL INFORMATION:

APPLICANT: DEMEREY, PHILIPPE

APPLICANT: AMORY, ANTOINE

TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH

PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESS: WILLIAM BRINKS HOFER GILSON &amp; LIONE, P.C.

STREET: 2000 K Street, N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-474,545

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadlano, Esq.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4.21-43

TELEPHONE: (202) 429-0625

TELEFAX: (202) 293-1850

TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 4464 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

US-08-474-545-9

Query Match 99.7%; Score 2786.4; DB 1; Length 4464;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatgggaacacgacacgacatctgtccactcttttggccctggtgtatatacaact 60  
|||||  
Db 1250 gatgggaacacgacacgacatctgtccactcttttggccctggtgtatatacaact 1309  
OY 61 tggagctctatgagatgtagcacaagaacgagagtgagagtgaaatcacaacacg 120  
|||||  
Db 1310 tggagctctatgagatgtagcacaagaacgagagtgagagtgaaatcacaacacg 1369  
OY 121 gctgactcttttgagagctgtgtcaagtgctgatalatccagaaacccaagtcagta 180  
|||||  
Db 1370 gctgactcttttgagagctgtgtcaagtgctgatalatccagaaacccaagtcagta 1429  
OY 181 atctacgttcgactcaagatgtgacccaagaatgtgagcgccgacccgctacataagttca 240  
|||||

Db 1430 attatcgtttcgacatcagattggacacaaagatgtgacggcgtgacgtatgatttga 1489  
OY 241 agcaaaagaaatgagtggtgtgtgtgagaaggaacacgcaaatctttataatgaaaa 300  
|||||  
Db 1490 agcaaaagaaatgagtggtgtgtgtgagaaggaacacgcaaatctttataatgaaaa 1549  
OY 301 gatgtcgaagatgacgtacaaacccgctgtgaagcaagctattatagatgttcaaacag 360  
|||||  
Db 1550 gatgtcgaagatgacgtacaaacccgctgtgaagcaagctattatagatgttcaaacag 1609  
OY 361 ggtcgtgtaaaacttaagccagccgtaactcttgggggaagannaagcggcttaaggt 420  
|||||  
Db 1610 ggtcgtgtaaaacttaagccagccgtaactcttgggggaagannaagcggcttaaggt 1669  
OY 421 catgacacacagcaaaatagatgatalccagtgacatctgtgagaagtcaagcttctgt 480  
|||||  
Db 1670 catgacacacagcaaaatagatgatalccagtgacatctgtgagaagtcaagcttctgt 1729  
OY 481 caagatgtaacccgtgttttgcaagtgacatcttcaacataatttggaggttccagttg 540  
|||||  
Db 1730 caagatgtaacccgtgttttgcaagtgacatcttcaacataatttggaggttccagttg 1789  
OY 541 gcaacctgataatcagatctatataaaagagtgactaacatctctatcaattctca 600  
|||||  
Db 1790 gcaacctgataatcagatctatataaaagagtgactaacatctctatcaattctca 1849  
OY 601 ggaagatcttcctgaaggaacacacaaatgaagtggttgaatgataagctgaaat 660  
|||||  
Db 1850 ggaagatcttcctgaaggaacacacaaatgaagtggttgaatgataagctgaaat 1909  
OY 661 ccgagtgtaacccatctgacacacataatgaagtggttgaatgataagctgaaat 720  
|||||  
Db 1910 ccgagtgtaacccatctgacacacataatgaagtggttgaatgataagctgaaat 1969  
OY 721 ttctgtaattcgtcgtcactcaatgacgtatgacacaaatgaatccttaagtcgagat 780  
|||||  
Db 1970 ttctgtaattcgtcgtcactcaatgacgtatgacacaaatgaatccttaagtcgagat 2029  
OY 781 ttcaagtagaagcgaggttgaacgagatctgtgacgttcaagtgaggaagatcca 840  
|||||  
Db 2030 ttcaagtagaagcgaggttgaacgagatctgtgacgttcaagtgaggaagatcca 2089  
OY 841 gatgtgagccatcctgttccattcaaacagatgagtgatcaagcgagtgatcaact 900  
|||||  
Db 2090 gatgtgagccatcctgttccattcaaacagatgagtgatcaagcgagtgatcaact 2149  
OY 901 cgttaatgcttaattcattcaacagatgagtgatcaagcgagtgatcaact 960  
|||||  
Db 2150 cgttaatgcttaattcattcaacagatgagtgatcaagcgagtgatcaact 2209  
OY 961 acacagaaagcaaacacattgaagctgtggacacaaactctcaagtaaatgttctt 1020  
|||||  
Db 2210 acacagaaagcaaacacattgaagctgtggacacaaactctcaagtaaatgttctt 2269  
OY 1021 cttaatgacagtgcaacgaggttctgttaacaaatcgttaccatgaagtgagtgag 1080  
|||||  
Db 2270 cttaatgacagtgcaacgaggttctgttaacaaatcgttaccatgaagtgagtgag 2229  
OY 1081 ggtgtgtggaagcaacggttcaacaaacttgaaatgtgtatcaatgagtgag 1140  
|||||  
Db 2330 ggtgtgtggaagcaacggttcaacaaacttgaaatgtgtatcaatgagtgag 2389  
OY 1141 acagagcaagcgcttaccgcaacgagctgtgtatccttgaactcgaacttgacacaaat 1200  
|||||  
Db 2390 acagagcaagcgcttaccgcaacgagctgtgtatccttgaactcgaacttgacacaaat 2449  
OY 1201 ggaagcagaggaatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260  
|||||  
Db 2450 ggaagcagaggaatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2509  
OY 1261 aaacatattagcccaagaatataagaatgagtgatcattcaatgagtgagtgag 1320  
|||||  
Db 2510 aaacatattagcccaagaatataagaatgagtgatcattcaatgagtgagtgag 2569



Query Match		99.7%	Score 2786.4	DB 1	Length 4464
Best Local Similarity		100.0%	Pred. No. 0		
Matches 2793		Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	gatgggaacacgaacagatcatctgtccacatttttgcctgtgtgtatatacaact 60			
Db	1250	GATGGGAACACGACAAACATCATTTGTCATTTTGGCCGCTGATATACCT 1309			
Qy	61	tggagctatgagtggtgccaagaacgagtggtgagtgatatacaacacg 120			
Db	1310	TGGAAGCTATGAGATGTGGCCAAAGACGAGGTGGGCTGAATACGATTCATACACCG 1369			
Qy	121	gctgacatcttctggagctgtgtgcaagtgctgatalctccaggaacccaagtcagga 180			
Db	1370	GCTGACTCTTTTGGAGCTGTTGCAAGCTGCTGATATTCACGGAACCCAAAGTCAGTAGGA 1429			
Qy	181	attatgcttcgacatcaagaatltgaccaagaatgtgagcgctgaccgctacatagatta 240			
Db	1430	ATTATCGTTCGCACTCAAGATTGGACCAAGATGTGAGCGCTGACCGCTACATGATTTTA 1489			
Qy	241	agcaagaagaatgagtggtgtgctgtgagaagaacagccaattttatataaanaa 300			
Db	1490	AGCAAGGAATGAGGTGTGGCTGTAGAGGAACAGCAAAATTTTATATGAAAAA 1549			
Qy	301	gagctgagagatgcagctaaacccgcgtgaaagcaagcttattagatgctcaaacag 360			
Db	1550	GATGCTGAGATGACAGCTAAACCCGCTGAACCAAGCTTATTGATGCTTCAAAACCG 1609			
Qy	361	gtgtgtgtttaaactagccagcgttlaactcttggggaaggnnaagcgctttagt 420			
Db	1610	GTGTGCTTAAACTTACGACGCGCTTAACCTTGGGGAAGNNNAAGCGCTTACGCTT 1669			
Qy	421	catgacgacacagcaataaagatacttcagtgacatctgtgaagaatgcaagtcctgt 480			
Db	1670	CATGACGACACGCAAAATGAGATTTCCAGTGACATCTGTGAAGATGCAAGTCTTGCT 1729			
Qy	481	caagaatgaacccgtgttctggcaagtaacctccaacatattcttggaggtctcgatlg 540			
Db	1730	CAAGATGTAACCGGTGTTTGGCAGGTACCTTCCAACATATTTTGGAGCTTCCGATG 1789			
Qy	541	gcacctgataatcacagtaacttataaaaaaagtgactaacatctctataacttca 600			
Db	1790	GCACCTGTATTAACAGATCTTATTAAMAAAGGTACATMACATCTCATTAATTTCTCA 1849			
Qy	601	ggaagctcttcctggaagaacactcaataaagtgagcttgaagatgagtgataat 660			
Db	1850	GGAGATCTTCTCGAAGAACTACCAATATTAAGTGCTTTTAAATGATAGCTGGAATAT 1909			
Qy	661	ccgagttacccaactcgaacaacatlaattaaagtccttcgagcggtgcaacgttca 720			
Db	1910	CCGAGTTACCCATCTGCAACACTTAATTTAACAGTCCCGCGGGGTGCAACGTCAT 1969			
Qy	721	tttctgtatattcgatccaactcaagcagtcataagcaacatlaatacttaatcgag 780			
Db	1970	TTTCTGTATATTCCTCCACTCATCATGAGTCTATGACACATTAATTAATCTAATGGGAT 2029			
Qy	781	ttacaagaagaaagcggttaaaacgagtcctcgagcggttacttaggggaagatca 840			
Db	2030	TTACAGTAGAAGAGGGGTTTAAAGGATCTGTGACGGTTACTTAAAGGAAGTTCA 2089			
Qy	841	gatgtgaacccatactctgtcaatcaaaacagatggatcaggaacgaagtgataact 900			
Db	2090	GATGTGAGCAATCTCTGTCCATTCACACAGATGGCTATACAGGAAGGAGTGAATCT 2149			
Qy	901	cgtaattgtctaatatcatcacagtaactatacaagaagatgacttgggaatactat 960			
Db	2150	CGTATGTGCTTAATTCACAGTACTATTCAGGAGATGATCTTGGGAATACCTAT 2209			
Qy	961	acacagaagacacaaacttaaatcttgggacccaactctaaccaagtaaatcttct 1020			
Db	2210	ACACGAAGAACACAACTTTAAAGTGTGGCAACCACTTCTACTCAAGTAATGTCTT 2269			
Qy	1021	cttlatgacagtgcacacgggtctctgtaacaaaatcgtacctaagcgcacacggccat 1080			
Db	2270	CTTATGACAGTGCAACGGGCTTCTGTAAACAAAATCGTACGTATGACGGCATCGGCCAT 2329			
Qy	1081	gggtgtgggaagcaacggttatacaaaacttgaaaatgtgtatatactatgtagta 1140			
Db	2330	GGTGTGGGAAGCAACGGTTATCAAAAACCTTGAAAATTTGATATACGTATGAGGTA 2389			
Qy	1141	acagaccaaaggtcttaacccgaacggctgtgtgacttccctatgaacgtgttcaaccaat 1200			
Db	2390	ACAGGCCAAGGCTCTACCCGAACGGCTGTGATCTTATGCAACTGCGATTCACCAAAAT 2449			
Qy	1201	ggaacgaagagcatgaltgtgacacgtgctgataaacagatcccgctgbcggaacagtat 1260			
Db	2450	GGAAGGAGGAGCATGATTGTGGACCTGGCTAAACGATCTGTGCGCTGGAACATGAT 2509			
Qy	1261	aaacatataagcccaagaatatagaagatgaggtcaactatgaaatggaatgtccgtgac 1320			
Db	2510	AAACATATTACGCCCAAGAAATATAGAAATGAGTGCATCTATGAAATGGATGTCCGTGAC 2569			
Qy	1321	tttccattgaccttaattccgggtatgaaaaataaagggaagatatttgctcttaagaa 1380			
Db	2570	TTTTCCATTGACCTTATTCGGGTATGAAAAATAAAGGAAGATTTGGCTTTACAGAA 2629			
Qy	1381	aaaggaacaaaggccctgacaaagtaaaagacgggatalagattcccttaaacacacttgg 1440			
Db	2630	AAAGGAACAAAGGCCCTGACAAACGTAAAGACGGGATAGATTCTTAAACAACTTTGG 2689			
Qy	1441	attaactatgttcaagcttatgcctgttctcgacatcaacagtgatgaaactgtatca 1500			
Db	2690	ATTACTCATGTTCAAGCTTATGCTTGGCTTTTGGCATTCMACAGTGTGATAAACTGATCA 2749			
Qy	1501	accgaagaataattggggtatgaaacccctgcaactagatgtcttcgaaaggacgtatgct 1560			
Db	2750	ACCCAAGATTAATTTGGGTTATGACCTCCCAACTATGATGTCTCTAAAGGCGATGATGCT 2809			
Qy	1561	acaatatgcgaatgtaatgctgcataaaagatttaaggaaatggtcttcttaaccat 1620			
Db	2810	ACAAATGCAATATGCTATGCTGATTAAMAAAGTTAAAGGAATGTTCTTTCACCTCAAT 2869			
Qy	1621	cgatgaacacattgggggttaacatggaatgttctataatacaacttgcacgcgaatc 1680			
Db	2870	CGTGAACACATTGGGGTTTAAATGATGTGTTGCTATTAATCAATACCTTCCACGCAAAATC 2929			
Qy	1681	tctgacctgataaatacttaccagaatattaccagatgagtgatccagtgatata 1740			
Db	2930	TCTGACTTGATTAATTTGACAGATATTTATACCGTACGATGATGCAAGGTAAATTA 2989			
Qy	1741	ccaacgatacaggtacgtgaatgaaatlgcangcngaaggccaatgglttcaaaaatt 1800			
Db	2990	CCAACGGAATCAGGTACTGGAATGAAATGCAAGCNGAAGGCCAATGGTTCAAAAATTT 3049			
Qy	1801	attattgattcccttaagaatgagtgcaatgagatcaataatgaaggcttccgtttgac 1860			
Db	3050	ATTATTGATTCCCTTAAGTATGGGTCAATGAGTATCAATATGACCGCTCCCTTTTGAC 3109			
Qy	1861	ttaatggcgctccttggaaaagacagatgtccaagcgccctcggaagttcatgtatt 1920			
Db	3110	TTAATGGCCCTGTGGAAAACAGATGTCCAAAGCTGCGCGAGCTTCATGCTATTT 3169			
Qy	1921	aatccaggaatgacattttagcgtgagcaatggaacgggtggaacccctgcaactgcagat 1980			
Db	3170	AATCCAGGAATTTGACATTTTACGCTGAGCCATGACGGGTGGAACCTCTGCACAGCCAGAT 3229			
Qy	1981	gataagcttcgtacaagaagagctcaaaaggaatgagagtgaggtgtttaaagcaat 2040			
Db	3230	GATCAGCTTCTGCAAAAGAGCTCAAAAAGGATGGGATGACGGGTGTTTATATGCAAT 3289			
Qy	2041	ttaacgaacggcttggacggaaatgctcttgatcttctcgcctcaaaagtttgggaaggt 2100			
Db	3290	TTACGAAGACGGCTTGGACCGCAATGTCTTTGATTTCTCCGCTCAACGTTTGGACAGGT 3349			
Qy	2101	gcaacagcttlaactlgaatgaatgagtggttggaggagatlaatgacttacc 2160			



|||||  
Db 3350 GCACACAGCTTAACTGATGAAATTAGAAATGGCTTAGGGGATATTAATGATTTACC 3409  
Oy 2161 tcttcacaggtgagacaattaaactatgtaacaaagtcataactaacacccttggac 2220  
Db 3410 TCTTCACAGAGTGAACAATTAACTATGTCAACAAGTATATACTACACCTTTGGGAC 3469  
Oy 2221 aaaaataagccataagcaaacctaaatgaltccgaagcgagtcggatlaaataatgatactc 2280  
Db 3470 AAAATACCCCTAAGCAATCTTAATGATTTCCGAAGCCGATCGATTAATAATGATGAATC 3529  
Oy 2281 gacaaagcagttgataagctcaagaggttccactatgcaagcgaggaagaaatg 2340  
Db 3530 GCACACAGCAATTGTTATGACTTCAACAGGCGTTCCATTCAAGCAGGCGGAAGAAATG 3589  
Oy 2341 ctctgtaaaaaagcgagcaacgaataagttataatgcaagcgagtcgagcaatgagttc 2400  
Db 3590 CTTGCTANMAAAGGCGGCAACGAAATAGTTATATGACAGCGCATCGGCTAATGATTT 3649  
Oy 2401 gattgagcgagaaagctcaatataatccagatglttcaactatataagcgagctaaaccac 2460  
Db 3650 GATTGAGAGCAAAAGCTCAATATCCAGATGTTTCAACTATATATACCGGCTAATTCAC 3709  
Oy 2461 ctctgcttcatcaacccagcttccgcatgacgacagcaatgaatcaatlaagcaccctc 2520  
Db 3710 CTTGCTTATATCAACCAAGCTTCCGATACGACACTAATGAATCAATAGCCACTTC 3769  
Oy 2521 caattcttaataatgctcagagaacacagtgagctatgaatlaactgaatcagttatacaaa 2580  
Db 3770 CAATTCCTAATAGTCCAGAGAACACAGTGGCTATGATTAATGATCAATGTTAATAA 3829  
Oy 2581 gacaaatggggaaaaatcatatgltgttataaaccataaactatgacaacataat 2640  
Db 3830 GACAAATGCGGAATATATCATTTGTTTATTAACCAATATAACTGATAGAACCATCAAT 3889  
Oy 2641 ttgcagagcgaggaatgggaaatcaatgctacagcgagtaaggaagatccaccctc 2700  
Db 3890 TTGCGAGCGGGAAATGGCAATCAATGCTACAGCCGTTAAGTAGAGATCCACCCTT 3949  
Oy 2701 ggtcaagcagagggaaagtgtcctcaagtaacagatataatcagtgatccctcacaagag 2760  
Db 3950 GGTCACAGCAGGAGAGTGTCCAAAGTACCAAGTATATCTATGATGATGCTTCAACAGAG 4009  
Oy 2761 gtaagccagacacagtgataaaagtaataagaa 2794  
Db 4010 GTAAAGCCACACACGCTAAAGTAATAGAAAA 4043  
RESULT 10  
US-08-478-341-9  
Sequence 9, Application US/08478341  
Patent No. 5817498  
GENERAL INFORMATION:  
APPLICANT: DEWEER, PHILIPPE  
INVENTOR: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILHJAN BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,341

FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhlem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-45  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-478-341-9  
Query Match 99.7%; Score 2786.4; DB 1; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 gatgggaacacgaaacagatcatgtccactatcttgcctctggtgattatcaact 60  
Db 1250 GATGGGAACACGAAACGATCATTTGCTCAATTTTGGCCCTGCTGATTAACAACCT 1309  
Oy 61 tgaagctatgagatgagcgcaaaaagcagaggtgagggctgatacgaattcaacacacg 120  
Db 1310 TGAAGCTATGATGTGCGCAAAAAGCAGAGGTGGCTGATAACGATTTCAATCAACCG 1369  
Oy 121 gctgactcttctgagagctgtctgaagtcgataatccagaagaaacccaaagtcagtaga 180  
Db 1370 GCTGACTCTTTTGGAGCTGTTCGAAGTGTGATATTCAGAGAAACCAACGAGTAGA 1429  
Oy 181 attatgcttcgacactcaagattggaacaaagatgagcgctgacgctacatagatt 240  
Db 1430 ATTATGCTTCGACACTCAAGATGAGATGACCAAGATGTGAGCCCTGACCCCTCATGATTTA 1489  
Oy 241 agcaaaagaaatgagtggtgctgtgtagaaggaacagcgaatcttataatgaaaaa 300  
Db 1490 AGCAAAAGAAATGAGTGTGTGCTTTGTGAAAGGAACAGCCAAATTTTATATGAAAAA 1549  
Oy 301 gatgtgaggaatgcaagctaaacccgctgtaagcaagcttattagaatgctcaaacag 360  
Db 1550 GATGCTGAGATGACGACTAAACCCGCTGTAAAGCAACGCTTATTAGATGCTTAACACAG 1609  
Oy 361 gttcgtgttaacttaagccgctgtaactcttggggaagmmnaagcggtcttaaggtc 420  
Db 1610 GTTCTGTTAAACTTAAGCCAGCCGCTTAACCTTGGGGAAGNNNAACGGCTTTACGGTT 1669  
Oy 421 catgacgacacagcaaaataagatataatccagtgacacatctgtagaagaatgaaagtcgtc 480  
Db 1670 CATGACGACACAGCAAAATTAAGATATCCAGTGAACATCTGGAAGATGCAAGCTTGGT 1729  
Oy 481 caagaatgaacgcgtcttggcaggtacatccacaatatcttggaggctccagattgg 540  
Db 1730 CAAGATGAACCGCTGTTTGGCAGGTACCTCCACATATTTTGGAGGTCGATTTGG 1789  
Oy 541 gacactatataatcaagtaacttataaaaaaggtgactcaacatcttcaatctca 600  
Db 1790 GACCTGATATACACAGTACTTTATTAATAAAAGCTGACTAACATCTCTATCAATTCCTCA 1849  
Oy 601 ggaagcttcctcgaaggaacatacacaataaagtgagcttcaaaatgataagtgaaat 660  
Db 1850 GGAAGATCTTCTGAAGAAATACCAATTAAGTGGCTTTAATGATAGCTGAATAT 1909  
Oy 661 ccgaatcaacatctgcaacaatataatlaacaggtccctcgcggtgtagacagctcat 720  
Db 1910 CCGAGTTAACCATCTGACAAATTAATTTAAAGTCCCTCCGCGGCTGACACAGCTACT 1969  
Oy 721 ttctgataatctcgtccactcaatgagctatgagacaatataatctctaagcgat 780



|||||  
Db 1970 TTTTCGTAATATTCGTCACATCATCATGTATGACACAATTAATTAATCTTAATCGGAT 2029  
Qy 781 ttacaagtaagaacgagggglltaaaacgagatctcgtaagcttaaggggaagatcca 840  
|||||  
Db 2030 TTACACAGTAGAAGGGGGGTTTAAACGATCTCGTAGCGGTACTACTGAGGGAATCCA 2089  
Qy 841 gatgtgagccatactctgtlccatltcaaacagatggtatcagggcaaacggtgaacct 900  
|||||  
Db 2090 GATGTGAGCCATTAATCTCTGTCATTCAAACAGATGSGCTATGAGGCAAGAGGTGATACCT 2149  
Qy 901 cytaatgtgttaataatctacacagtaactacattcaagagatgatactctgggaataacct 960  
|||||  
Db 2150 CGTAATGTGCTTAATTAATTCACACAGTACTACTATTCAGAGAGATCTTGGGAAATCTAT 2209  
Qy 961 acacagaagaacaaacactttaaagltctgggacacacacttctaaccaagtaatttctt 1020  
|||||  
Db 2210 ACACAGAAAGCAACACTTTTAAAGTCTGGGACCACTTCTACTCAAGTAAATGTTCTT 2269  
Qy 1021 ctctatgacagtgcaacggtctgttcaacaaaatcgtactatgaagcgatcgagccat 1080  
|||||  
Db 2270 CTTTATGACAGTGCNAACGGGTTCTGTAAACAAAATCGTACTATGACGGCATCGGGCCAT 2329  
Qy 1081 ggtgtgtgtggaagcaacggttaatacaaaaccttgaaaattggtattacaatgtataggtta 1140  
|||||  
Db 2330 GGTGTGTGGGAAGCAACGGTTAATCAAAACCTTGGAATTTGGTATTACATGTATGAGTA 2389  
Qy 1141 acaagacaaagctctaacacgaacggtctgttgaaccttcttgcaacgtcgatgacaaat 1200  
|||||  
Db 2390 ACAGGCCAAGGCTTACCCGGAACGGCTGTGATCTTATGCAACGCTGCGATTCACCAAAAT 2449  
Qy 1201 ggaacgagaagcatgatatgttgaacctgtgcaaaaacagatccctgtgtgtgtgacaggtgat 1260  
|||||  
Db 2450 GGAAGCAGAGCATGATGTGTGAGCCTGTGCTAAACAGATCCTGCTGGCTGGAACAGTGAT 2509  
Qy 1261 aaacatataagccaaaagaatatagaagatgagtgatccatgaaatggaatgtccgtgac 1320  
|||||  
Db 2510 AAACATATTAGCCCAAAAGATATAGAGATGAGTCAATCAAGAAATGATCTCCGTGAC 2569  
Qy 1321 ttttccattgaccccttaattcgtgtgtatgaaaaataaagggaatatttgctcttcaaga 1380  
|||||  
Db 2570 TTTTCCATTGACCTTAATTTCCGGGTATGAAAATAAAGGAATATTTGGCTTTACAGTA 2629  
Qy 1381 aaaggaacaagggccctgacaacgtaaaagcgggagatagatcccttaaaaacacttgg 1440  
|||||  
Db 2630 AAAGGAACAAGGGCCCTGACAAAGCTAAAGAGGGGATGATTCCTTAAACAACTTGCG 2689  
Qy 1441 attactcatgtcagcttaagcttgcgttctcgcaatcaacagtgctgatagaacgatacca 1500  
|||||  
Db 2690 ATTACTCATGTTCACCTTATGCTGTTTTCCGATCTAACAGTGTGATGAAACTGATCCA 2749  
Qy 1501 acccaagaataattgggtgtatgaacctcgcaactatgatatctctgaaagggaagtatgct 1560  
|||||  
Db 2750 ACCCAAGATTAATGGGGTTATGACCTCGCAACTATGATGTTCTCGTAAGGGGAGATGCT 2809  
Qy 1561 acaaatgtcgaatgtgtaatgctgtatlaaaagagttlaaggaatgtgttcttctactccat 1620  
|||||  
Db 2810 ACAATGTGGAATGTAAATGCTGTATTAAGAGAGTTTAAGGAATGTTCTTCCACTCCAT 2869  
Qy 1621 cgtgaagaacattgggtgttaacatgtgagtgtgtctataatacaataccttggcagcaaac 1680  
|||||  
Db 2870 CGTGAACACATTGGGGTTAACGTGATGTGCTATTAATCAATACCTTTGGCCACGCAAAATC 2929  
Qy 1681 tctgactcgaataaattgttacaagaatactatctacgtaacgataccaagtaattata 1740  
|||||  
Db 2930 TCTGACTTGATTAATATGTACACAGATATTTATACCGTACAGATGATGAGGTATTTATA 2989  
Qy 1741 ccaacggaatcaggtacgtgaaatgaaatgcangcngaaaggccaaatgtgttcaaaaattt 1800  
|||||  
Db 2990 CCACAGGATACAGGTACTGGAATGAAATTTGCANGCNGAAAGGCCAATGTTCAAAAATTT 3049  
Qy 1801 attatgttcccttaagatattgggtcaatgtgataatataatgaaggtctcggttttgc 1860  
|||||

Db 3050 ATTATGATTCCTTAAGTATTGGTCAATAGATATCATATTGACCGCTTCGTTTGAC 3109  
Qy 1861 ttaatgagcgtctgtgaaagaacacagatggtccaaagctgctcgagcttcatgctatt 1920  
|||||  
Db 3110 TTAATGGCCCTCTTGGAAGAACACAGATGTCCAAAGCTGCTCGAGCTTATGTATT 3169  
Qy 1921 aatcagaagaattgcacttaccgtgtgagccaatggaacggtgtgnaacctctgcactgcagat 1980  
|||||  
Db 3170 AATCCAGGAATTTGACATTTACGGTGAAGCATGAGAGGGGGAAGCTCTGCACTGCGAGAT 3229  
Qy 1981 gatcagctcttgcaaaaaggagatcaaaaaggatggaagatgaggtgtgttataagaat 2040  
|||||  
Db 3230 GATCAGCTTCTGACAAAAGAGACTCAAAAAGCATGGGAGTGGCGTGTATATACAAAT 3289  
Qy 2041 ttacgaacaggttgagacgaatgcttctgtatcttctccgcacaaagtttgcgaaggt 2100  
|||||  
Db 3290 TTAGCAAGCGCTTGAGCGGCAATGTCTTGTATTCTTCGCTCAAGGTTTTCGACAGGT 3349  
Qy 2101 gcaacaggttactatgataatgaagaatgagcgttgaggggagatataatgacttacc 2160  
|||||  
Db 3350 GCACAGGCTTAAGTATGATCAATTTAAGATGGCGTTGAGGGAGATTAATGACTTTACC 3409  
Qy 2161 tcttcaacaggtgagaacaaatatactatgtcacaagtcataataactaaccttggagc 2220  
|||||  
Db 3410 TCTTACCGAGTGAGACAAATTAATCATGTGCACAGTCATGATTAACCTTTGGGAC 3469  
Qy 2221 aaaaatagcccttaagcaatcttaattgatttccgaagcgagatcgaattaaatgataac 2280  
|||||  
Db 3470 AAAATAGCCCTTAAGCAATCTTAATGATTTCCGAAAGGATCGGAATTAATGATGAATC 3529  
Qy 2281 gcaacagcagttgttatagacccacacagcggttccatltcatgcaagcgagggaagaatg 2340  
|||||  
Db 3530 GCACAGCGGTGTTATGACCTCACAGCGGTTCCATTCATGCAAGGGGGAAGAAATG 3589  
Qy 2341 ctctgtanaaaaggcggcaagcaacatagttataatgacggcagatggttcaatgaatt 2400  
|||||  
Db 3590 CTTTGTANAAGAGGGGCAACACATAGTATTAATGACAGGGGATGCCGTCAATGAGTT 3649  
Qy 2401 gattgagcagaagaagctcaatataatcagaatgttccaactatataggggcttaaccac 2460  
|||||  
Db 3650 GATTGAGCAGGAAGCTTAATATCCAGATGTTTTCATTAATTAAGCGGCTAATCCAC 3709  
Qy 2461 ctctgtcttcatcacccagcccttccgcatgacgaacagcttaatgaaatcaataaccac 2520  
|||||  
Db 3710 CTTGCTGTGATCACCCACCTTCGSCATGACGACGATTAATGAATCAATGCCACCTC 3769  
Qy 2521 caattccataatagttccagaagaacacagtgccctatgaattcaatgataatgataa 2580  
|||||  
Db 3770 CAATTCCTTAATTAATGTCACAGAAACACAGTGGCTATGAATTAATCATCATGTTATAA 3829  
Qy 2581 gacaaatggggaatatataatgttgttlaaaacccaaataaaaactgttagcaacctcaat 2640  
|||||  
Db 3830 GACAAATGGGGAATATCTATTGTTTATTAACCCAAATTAATTAATTAATTAATTAAT 3889  
Qy 2641 ttgcccagcgggaaatgaggcaatcaatgctacagcggtaaggtlaaggaatccaccc 2700  
|||||  
Db 3890 TTGCGAGGGGGAATATGGCAATCAATGCTACGAGCGGTAAGTAGAGAAATCCACCTT 3949  
Qy 2701 ggtcaagcagaaggaggttccaagtaacaggtatataatcaatgataatccttcaagaag 2760  
|||||  
Db 3950 GGTCAAGCAGAGGAGAGTGTCCAAAGTACAGGTATATATATATATATATATATAT 4009  
Qy 2761 gtaagcccaagacaggtlaaaaaagtaaatgaa 2794  
|||||  
Db 4010 GTAAAGCCAGACCGTAAAGTAATGAAAA 4043  
|||||

RESULT 11  
US-08-96-733-8  
; Sequence 8, Application us/08996733  
; Patent No. 6074854  
; GENERAL INFORMATION:  
; APPLICANT: DEMEER, PHILIPPE

APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 925 Page Mill Road  
CITY: Palo Alto,  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,733  
FILING DATE: 23-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,293  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/174,893  
FILING DATE: 28-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: BE 09301278  
FILING DATE: 19-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: BE 09300744  
FILING DATE: 15-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: BE 09201156  
FILING DATE: 28-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Debra J. Glaister, Esq.  
REGISTRATION NUMBER: 33,888  
REFERENCE/DOCKET NUMBER: GC446C1-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 846-7620  
TELEFAX: (650) 845-6504  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-996-733-8

Query Match 99.7%; Score 2786.4; DB 3; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatgggaacacgacacatcattgtccactattttgacctgtggtgtaatacaacct 60  
DB 1250 gatgggaacacgacacatcattgtccactattttgacctgtggtgtaatacaacct 1309

QY 61 tggagttatagatgtgtgacaaagaaggggtgtgggtctgaatacagttccaatacagc 120  
DB 1310 tggagttatagatgtgtgacaaagaaggggtgtgggtctgaatacagttccaatacagc 1369

QY 121 gctgaactcttggagctgtctgcaagtgctgatatctcagaagaacccaagtcagtaga 180  
DB 1370 gctgaactcttggagctgtctgcaagtgctgatatctcagaagaacccaagtcagtaga 1429

QY 181 atatacgttcgaactcaagaatggacaagaagatgtgagcgttgacgctacatagattta 240  
DB 1430 atatacgttcgaactcaagaatggacaagaagatgtgagcgttgacgctacatagattta 1489

QY 241 agcaaaagaaatgaggtgtgtgctgtgtagaagaacagccaattttcttaataatgaaaaa 300

|||||  
DB 1490 AGCAAAAGAAATGAGGTGTGGCTTGTAGAAAGAAACAGCCAAATTTTATTAATGAAAAA 1549  
QY 301 gatgtgagatgcagtgtaaacccgcgtgtaagcaacctatatagtgcttaaacacag 360  
DB 1550 gatgtgagatgcagtgtaaacccgcgtgtaagcaacctatatagtgcttaaacacag 1609  
QY 361 gtgtgtttaaacttaagccagcgtttaacctctgtggggaagmnnaagcagcttaagtc 420  
DB 1610 gtgtgtttaaacttaagccagcgtttaacctctgtggggaagmnnaagcagcttaagtc 1669  
QY 421 gatgagacacagcaataaagatatccagtgacatctgtgaagaatgcaggtcttggc 480  
DB 1670 gatgagacacagcaataaagatatccagtgacatctgtgaagaatgcaggtcttggc 1729  
QY 481 caagatgtaacgcgtgtttggcagagtaacctccaacatcttlttgaggcttcagattg 540  
DB 1730 caagatgtaacgcgtgtttggcagagtaacctccaacatcttlttgaggcttcagattg 1789  
QY 541 gcaactgataatcacagttacttattaaaaaagggtgaatacaatctataatctca 600  
DB 1790 gcaactgataatcacagttacttattaaaaaagggtgaatacaatctataatctca 1849  
QY 601 gggagatcttccgtgaagaagaactccaataaagaatgtgtcttaaatgataatgataat 660  
DB 1850 gggagatcttccgtgaagaagaactccaataaagaatgtgtcttaaatgataatgataat 1909  
QY 661 ccagattaccatcttgacaacatlaattaacagttccctgcgaggtgacacagttact 720  
DB 1910 ccagattaccatcttgacaacatlaattaacagttccctgcgaggtgacacagttact 1969  
QY 721 ttctgtataatccgtccactcatgacgtctatgacacaataataatccaaatgtcgat 780  
DB 1970 ttctgtatattccgtccactcatgacgtctatgacacaataataatccaaatgtcgat 2029  
QY 781 ttcaaatgaagaagggtgttaaaacggtctgtgagcgtttactcagggggaagatcca 840  
DB 2030 ttcaaatgaagaagggtgttaaaacggtctgtgagcgtttactcagggggaagatcca 2089  
QY 841 gatgtgacacatctctgtccatctcaaacagatlgagctacaagcaagcaggtgatacct 900  
DB 2090 gatgtgacacatctctgtccatctcaaacagatlgagctacaagcaagcaggtgatacct 2149  
QY 901 cgttaatgtcttaataatcaacagttactatctagagagatgactctgggaatacctat 960  
DB 2150 cgttaatgtcttaataatcaacagttactatctagagagatgactctgggaatacctat 2209  
QY 961 acaacagaagaacacacacttaaaagtctgggcaaccaactctacagaagaatgtctct 1020  
DB 2210 acaacagaagaacacacacttaaaagtctgggcaaccaactctacagaagaatgtctct 2269  
QY 1021 ctctatagacgtgcaacgggtctctgtlaacaaaaatctgtacataagcagcatcgagccat 1080  
DB 2270 ctctatagacgtgcaacgggtctctgtlaacaaaaatctgtacataagcagcatcgagccat 2329  
QY 1081 ggtgtgtggaagaacaggttaatacaaaccttgaaaaaattgtgatacatgtatgagta 1140  
DB 2330 ggtgtgtggaagaacaggttaatacaaaccttgaaaaaattgtgatacatgtatgagta 2389  
QY 1141 acagggccaaaggtcttaacccgaagcgtgtgtgtactcttatgtcaactgtgatatccaat 1200  
DB 2390 acagggccaaaggtcttaacccgaagcgtgtgtgtactcttatgtcaactgtgatatccaat 2449  
QY 1201 ggaacgaagagcatgattgtggaacctggcctcaaaaacagatctctgctggaagaatgtat 1260  
DB 2450 ggaacgaagagcatgattgtggaacctggcctcaaaaacagatctctgctggaagaatgtat 2509  
QY 1261 aaacataatcagccaaagaatataagaagaatgagttcataatgaatgtcatgtccgtgac 1320  
DB 2510 aaacataatcagccaaagaatataagaagaatgagttcataatgaatgtcatgtccgtgac 2569  
QY 1321 ttctccatgacccataatctgggtgatgaaaaataaagggaagtatttggctcttaacaaa 1380  
|||||

Db 2570 TTTTCATTGACCTTAATTCGGGTATGAAAAATAAGGAAGTATTTGGCTTTACAGAA 2629  
 Qy 1381 aaaggaacaaagggccctgacacgtaagaagcgggatagatctccctaaacacttggg 1440  
 Db 2630 AAAGGAACAAAGGGCCCTGACACGTAAGAGCGGGATAGATTCCCTTAAACAACTTTGGG 2689  
 Qy 1441 attactcaltgctgacttgcctgttctgcacatcaacagtlgcaltgaaactgaccca 1500  
 Db 2690 ATTACTCATGTTCAGCTTATCCCTGTTTCCCATCTTAACAGTGTGAGAAACTGATCCA 2749  
 Qy 1501 acccaagataatttgggtgtatgacccctgcacactatgatttccctgaaggcgatctct 1560  
 Db 2750 ACCCAAGATAATTGGGGTGTATGACCCCTGCAACTGTGATGTTCCGAAAGGCGATATCT 2809  
 Qy 1561 acaaatcgaaatgtgtaatgctgcgtataaagaagttlaagaaatgttcttccactcat 1620  
 Db 2810 ACAATGCGATGTGAATGCTGCTATTAACAGCTTAAAGAAATGCTTCTTCACTCCAT 2869  
 Qy 1621 cgtgaacacaltgggggttaacatggaatgttgcataaactaccccttgcacgcaatc 1680  
 Db 2870 CGTGAACACATTTGGGGTGTAAACATGATGTTGCTATTAATCAATACCTTCCACGCAAAATC 2929  
 Qy 1681 tctgactcgataaataatgtgtccagaataatataatccgtaacgtatgacgtatata 1740  
 Db 2930 TCTGACTTGTGAATAATGTGTACAGATATTTATTCCTGATGATGATCGAGTAAATATA 2989  
 Qy 1741 ccaacgagatcagctgactgtaaatgaatgtcangcnaaagccaatgttcaaaaattc 1800  
 Db 2990 CCAACGAGTACGATGATCTGGAATGAATTTGCANCGAAAGCCCAATGTTCAAAATTT 3049  
 Qy 1801 attattgattcccttaaglatltgggtcaatgaatgatalcataltgacgcttccgtttgac 1860  
 Db 3050 ATTATTGATTTCCCTTAATGATTTGGGTCAATGATGATTAATGACGCTTCCGTTTAC 3109  
 Qy 1861 ttaatggcgctgcttggaaagacagatgtccaaagctgctgacgtatcatctatt 1920  
 Db 3110 TTAATGGCGCTGTGAAAAACACGATGTCCAAAGCTGCTCGAGCTTATCTAT 3169  
 Qy 1921 aatccagaatgtcacttgaatgagtcagtcacatgagcggtggaaactctgcactgcaat 1980  
 Db 3170 AATCCAGAAATTTGACCTTTAGCGTACGATGAGCGGTGGAACCTTGCATGCGACAT 3229  
 Qy 1981 gatcagcttctgacaaaagagcttcaaaaagcgatggagtagcggtttaaagcaat 2040  
 Db 3230 GATCAGCTTCTGACAAAAGAGCTCAAAAAGCATGAGGAGTGTATGACAT 3289  
 Qy 2041 ttaagaagcgcttggagaggaatgtcttctgattcttccgctcaaggttttgcagaagt 2100  
 Db 3290 TTACGAAGCGGTTGGAGCGCAATGTCTTTGATTCTCCGCTCAAGGTTTGGACAGGT 3349  
 Qy 2101 gcaacaggtcttaactgataaagaatgagtcgttgaaggagatlaaacttaactacc 2160  
 Db 3350 GCAACAGGCTTAAGTATGATGCAATTAAGATGGCGTTGAGGGAGTATTAATGACTTTAAC 3409  
 Qy 2161 tcttcaacaggttgaacaatlaactatgtcacaaagtcataactacaccccttggagc 2220  
 Db 3410 TCTTCAACAGGTGAGACATTAATGCTCAACAGTCAATGATTAACCTTTGGGAC 3469  
 Qy 2221 aaatagagccctaaagaatccctaaatgattctcgaagcgagtcgataaaatgaataac 2280  
 Db 3470 AAATAGCCCTTAAGCAATCTTAATGATTCGAGGCGAGTGGATTAATTAATGATTAAC 3529  
 Qy 2281 gcaacagagcttgaatgactcacaagcgcttccatcaltgacaagcgagggaagaatg 2340  
 Db 3530 GCACAGAGCATTTGTTATGACCGTCAAGCGCTTCATTCATGCAAGCGGGGCAAGAAAG 3589  
 Qy 2341 ctctgtanaaaagcgagcaacagcaatagltataatgcaagcgatgaggttcaatgagttc 2400  
 Db 3590 CTCTGTANAAAAAGCGGACACCAATAGTTATATGACAGCGGATGCGGTCAATGATTT 3649  
 Qy 2401 gattgagcaggaagcacaataccagatgttccaaactatlatagcgaggtaactccac 2460  
 Db 3650 GATTGAGCAGGAAGCTCAATATTCACATGTTTCAACTATTATAGCGGGCTAATCCAC 3709

Qy 2461 ctctgcttgaatcaccagccttccgcatgacagacgtlaaataatgaaatagccactc 2520  
 Db 3710 CTCTGCTTTGATCACCACGACCTTCCGATGAGACAGGTAATGAATCAATGATCCACCTC 3769  
 Qy 2521 caatccctaataatgctcagaagaacagtgagcctatgaatlaactatgataatgaa 2580  
 Db 3770 CAATTCCTTAATATGCTCAGAGAACACAGTGGCTTGAATTAACGATCATGTATATAA 3829  
 Qy 2581 gacaaatggggaataatcatgttcttataaccccaataaactgtgcaacacatcat 2640  
 Db 3830 GACNAATGGGAATATCATTTGTTTATATACCCAAATTAACGTATGCAACCAATCAT 3889  
 Qy 2641 ttgcccagcgaggaatgggcaatcaltgactgacagcggtlaagtaggaatccaccct 2700  
 Db 3890 TTGCGACGCGGAATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3949  
 Qy 2701 gttcaagcagaaggaatgttccaaatgacaggtatataatgataatgataatgata 2760  
 Db 3950 GGTCAAGCAGAGGGAAGTGTCCAGTACAGGATATATATATATATATATATATATAT 4009  
 Qy 2761 gtaagccagacacaggttaaaatgaatgaana 2794  
 Db 4010 GTAACCCAGCACCGTAATAAATGTAATGANA 4043  
 RESULT 12  
 US-08-996-733-9  
 : Sequence 9, Application US/08996733  
 : Patent No. 6074854  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: DEMER, PHILIPPE  
 : APPLICANT: AMORY, ANTOINE  
 : TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
 : TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 : NUMBER OF SEQUENCES: 15  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genencor International, Inc.  
 : STREET: 925 Page Mill Road  
 : CITY: Palo Alto,  
 : STATE: CA  
 : COUNTRY: U.S.A.  
 : ZIP: 94304  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/996,733  
 : FILING DATE: 23-DEC-1997  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/472,293  
 : FILING DATE: 07-JUN-1995  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/174,893  
 : FILING DATE: 28-DEC-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09301278  
 : FILING DATE: 19-NOV-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09300744  
 : FILING DATE: 15-JUL-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09201156  
 : FILING DATE: 28-DEC-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Debra J. Glaister, Esq.  
 : REGISTRATION NUMBER: 33,888  
 : REFERENCE/DOCKET NUMBER: GC446c1-US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (650) 846-7620

TELEFAX: (650) 845-6504  
: INFORMATION FOR SEQ ID NO: 9:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 4464 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: genomic DNA  
: US-08-996-733-9

Query Match 99.7%; Score 2786.4; DB 3; Length 4464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2793; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggagacacgaacacgacatcatttgccttgcctgctgagtgatcaacct 60  
DB 1250 GATGGAGACACGACGACATCTCTCCTACTATTTTGGCCCTGCTGATTAACAACCT 1309  
QY 61 tggagctatgagatgagcgaagagcgaagtgaggctgaatacgaattccaatcaacg 120  
DB 1310 TGGAGCTATGAGATGAGCCAAAGAGGAGTGCGGCTGAATTCGATTCACACCG 1369  
QY 121 gctgacctttgagagctgttgcgaagtgctgatatccagaagaaacgaagtcagtaga 180  
DB 1370 GCTGACCTTTTGGAGCTGTGTGCAAGTCTGATATTCAGAGAAACCCACGTCAGTAGGA 1429  
QY 181 attatccttcgacatccaaagattgagccaaagatgagcgcctgacgcctacatagatcta 240  
DB 1430 ATTATCTTCGCACTCAACGATGAGACCAAGATGTGAGCCCTGACCCCTACATAGATTTA 1489  
QY 241 agcaaaagaaatgagatgctgctgtctgtagaagaaacagcaaatcttctaataagaaaa 300  
DB 1490 AGCAAAAGAAATGAGGTGCTGTGTGAAGGAAACGCCAAATTTTATATGAAAAA 1549  
QY 301 gatgctgagagatgagctaaacccgcgtgaagcagcctaataggatgcttcaaacag 360  
DB 1550 GATGCTGAGATGACGATTAACCCGCTGTAAAGCAACGCTTATTTAGTGTCTTAAACACAG 1609  
QY 361 gctgctgctttaaacttaacgacgccttlaactccttggggaagannaaacgcgcttaccgct 420  
DB 1610 GTCGCTGTTAACTTACCCAGCCGCTTAACCTTTGGCGAAGNNNAACGGCTTTACGCTT 1669  
QY 421 cgtgagacacacgaacaaataagatatccagtgacatcctgtgaagatgcaagtcctgct 480  
DB 1670 CATGACACACACGCAATAAGATATTCAGTGACATCTGTGAAGATGCAAGTCTTGCT 1729  
QY 481 caaagatgtaacccgctgttctggcaggtgaccttccaaacatacttttggaggcttccgattgg 540  
DB 1730 CAAGATGTAACCCGCTGTTTTGGCAGTACCTTCCACATATTTTGGAGGTTCCGATTGG 1789  
QY 541 gcaacatgataacacagtaacttlaataaaaggtagaacaactctctlaacatltca 600  
DB 1790 GCAACATGATATCAGCTACTTTATTAATAAAGCTGACTAACAATCTCTATCAATCTTCA 1849  
QY 601 gggagatcttccctgaagaaacacacacataaagtgagctttaaagttaactggaataat 660  
DB 1850 GGAAGATCTTCTGAAGAAACATACCAATATTAATGTGCTTAAATGATCTGCAATAT 1909  
QY 661 ccgagttaccacatctgacacacataaatttaacagctccctccgcgcgtgacacagctact 720  
DB 1910 CCGAGTTACCATCTTGACACATTAATTTAACACTCCCTCCGCGGTGACACAGCTACT 1969  
QY 721 tcttcgtatattccgtccactcagtcagtcctlaagacaaatlaataatctcaatgagat 780  
DB 1970 TTTTCGATATTCGCTCCACATGACGCTATACACAAATTAATATTCGTAATGCGGAT 2029  
QY 781 ttcaagatgaagagcggggtttaaagcgatcctcgtgacggcttactcctgaagggaagatcca 840  
DB 2030 TTCAAGATGAAGAGCGGGTTAAAGCGATCTCTGACGGTTACTCTGAGGGAAGATCCA 2089  
QY 841 gatgtgagcatactctgttccatctcaaacagatgagctactgaagcaagcagggtgatacct 900

DB 2090 GATGTGACCATACCTCTGTCATTCAAACAGATGCTATCAGCCAAACAGCTGATACCT 2149  
QY 901 cgtaatgtcttaataatcacaacagtaactatctcaagaagatgactcttggaaataactat 960  
DB 2150 CGTAATGTGCTTAATTCATACAGTACTACTATTCAAGAAATGATGCTGGGAATACCTAT 2209  
QY 961 acacagaagaacaaacactttaaagcttggcacaacactctcaactcaagtaaatgtctct 1020  
DB 2210 ACACAGAAGAACCAACACTTAAAGTCTGGGACCAACTCTCACTCAAGTAATGTTCCT 2269  
QY 1021 cttaataagatgcaagggcttctgttaacaaaatctgaccttgcagcagatggccat 1080  
DB 2270 CTTTATGACAGTGCACAGCGGCTTCTGTACAAAATGCTACTATGACGCGATCGGCCAT 2329  
QY 1081 ggtgtgtggaagaaacagtgatlaaacaacacttgaataatgattalacatgataagta 1140  
DB 2330 GGTGTGTGGGAAGCAAGGTTAATCAAAACCTGMAAATTTGATTAATCATGTATGAGGTA 2389  
QY 1141 acagccaaaggctctacccgacagcgtctgtatccttaagcaactgcaactgacaaat 1200  
DB 2390 ACAGGCCAAGGCTCTACCCCAACGCGCTGTGATCCTTATCAACTGCGGATTCACCAAT 2449  
QY 1201 ggaacgaagagcatgatttctggaacctgagctaaacagatccctctgagcttgaacagtagat 1260  
DB 2450 GGAACGAGAGCATGATTGTGGACCTGGCTTAAACAGATCTCTGCTGGAACAGTGAT 2509  
QY 1261 aaacatactacgcacaaagaaatatagaagctgaagtcatactalaagatgatactcgtgac 1320  
DB 2510 AAACATTTTACGCCCAAMGAATATGAAGATGAGGTCTATGTAATATGATGATGCCGTGAC 2569  
QY 1321 tcttccatgagccctaaactcggggtatgaaaaataaaggaagatattgctcttaacgaa 1380  
DB 2570 TTTTCCATTTACCTTAATTCGGGTATGMAAAATTAAGGGAAGTATTTGGCTCTTACGAA 2629  
QY 1381 aagaagaaacgaagggcccttgaacaaagctaaagcggggtataatccttaaaacaaacttgg 1440  
DB 2630 AAGGAACAAAGGGCCCTGCAACGTAAGACGGGGATTAATTCCTTAAACAACTTGGG 2689  
QY 1441 attactcatgttcagcttcaagctgtcttctgcacatcaacagctgcgaatgaaactgacaa 1500  
DB 2690 ATTACTCATGTTACGCTTATGCGCTTTCGATCTACATGATGATGATGATGATGATGATGAT 2749  
QY 1501 acccaagataatgtgggtcttgaacctcgcacactatgatttccctgaaggcagatgct 1560  
DB 2750 ACCCAAGATAATGTGGGTTTGAACCCGCACTATGATTTCTCTGAAGGGCAGTATGCT 2809  
QY 1561 acaaatgcgaatgtaatgtctcgtlaaagaagctttaaagaaatgagcttcttaacatccat 1620  
DB 2810 ACAAAATGCAATGTAATGCTCTGATTAAGAGATGTTAAGGAATGTTCTTTCACTCCAT 2869  
QY 1621 cgtgaacacatctggggttacaatgagatgtctctataatcaatcccttgcacgcaaatc 1680  
DB 2870 CGTGAACACATTTGGGGTTATACATGATGATGTTGCTATTAATCATACCTTGGCAGCAATTC 2929  
QY 1681 tctgactctgataaataatttgaacgaataatatttaccgttcagatgataccagtttaata 1740  
DB 2930 TCTGACTTGATTAATAATTTGACCAAGATTAATTAACGTCGATGATGACGGAATTTATA 2989  
QY 1741 ccaacgaatcagtaactggaagaaatgacangcgaagcgaatgagttcaaaaaatt 1800  
DB 2990 CCAACGATCAGGTACTGGAATGAATTAATTCANCKAAAGCCCAATGTTCAAAAATTT 3049  
QY 1801 attatgattcccttlaagatattgggtcaatgaglatcataattgaaggtcttcgtttgac 1860  
DB 3050 ATTATGATTCCTTAAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3109  
QY 1861 ttaatgagctgcttctgaaagaaacagatgtccaagctgctccgagcttcaatgctatc 1920  
DB 3110 TTAATGCGCTGCTTGGAAAAAGACAGATGTCCAAACCTCCTCGAGCTTCAATGCTAAT 3169  
QY 1921 aatccagaagattgcaactttagcgttgaagcattgaagcgggtgaagacatctgacatgacgat 1980  
DB 3170 AATCCAGGAATTCACCTTTTACGGTGAACCATGACAGGGGTGAACCTTGCACCTGCGAGAT 3229

```

QY 1981 gatcagcttctgacaaaaggagctcaaaaaggcaltggagtagcgglytlttaagacaat 2040
D 3330 GATCAGCTTCTGACAAAAGAGCTCAAAAAGCATGGAGTAGCGGTCTTAATGACAAAT 3289
QY 2041 ttacgaaacggttggaggaaggaatgcttcttcttccgctcaaggcttttggcgaagt 2100
D 3290 TTACGAAACGGCTTGGAGGGAATGCTTCTTAATCTTCCGCTCAAGGTTTGGCGACAGT 3349
QY 2101 gcaacagcttaactgatgatcaattaagaatgagcttggaggaggaattacgaacttacc 2160
D 3350 GCAACAGCTTAAGTGAATGAGATGAGATGAGGAGGAGATTAATGACTTTAC 3409
QY 2161 tcttcaccaggtagaacaattaactatgltcaaaagltcaglaaactacaccccttggagc 2220
D 3410 TCTTCACACAGCTGAGACAAATTAATGATGTCAAACTCATGATTAACACCTTGGGAC 3469
QY 2221 aaataagccctaaagcaatcccaatgatgtccgaagagagtcggatlaaataagtagaactc 2280
D 3470 AAAATAGCCCTTAGCAATCTTAATGATTCGGAAGCGATCGGATTAATGATGATGACTC 3529
QY 2281 gcaacaagagtgatgtatgacctcaaaagcgcttccatcatgcaagcgcggaagaatg 2340
D 3530 GCACAAGAGTGTATTATACCTCACAAAGCGCTTCATTCATGCAAGGCGGGAAGAAATG 3589
QY 2341 ctctgtanaaaaaggcggaacagacaatagltataatgcaagcgatgcgltcaatgagtt 2400
D 3590 CTTCGTAAAAAAGCGGCAACAGCAATAGTATTATATGACGCGATGCGTCAATGAGTWT 3649
QY 2401 gatggagcaggaagaagctcaatataccagatgtttcaactatlaagcgggtaataccac 2460
D 3650 GATTGGAGCAGGAAGCTCAATATCTCAGATGTTCACACTATTATAGCGGCTTAATTCAC 3709
QY 2461 ctctgtcttgtaacccagcctccgcatgacagcagcagtaataataatgacacccctc 2520
D 3710 CTTCGTCTTGTATCAACCCAGCTTCGCGATGACAGACGTAATGAATATATAGCCACTC 3769
QY 2521 caatctctaataagltccagagaacacagtgagcctatgaatlaactgacatgltataaa 2580
D 3770 CAATCTCTTAATAGTCCAGAGAACACAGTAGCGCTATGATTAACATGATCATGTATATAA 3829
QY 2581 gacaaatggggaataatacatgtgttataataccaataaaacgtgtgcaacatcaat 2640
D 3830 GACAAATGGGAATATATCTGTGTATTAACCCAAATTAACGTCGCAACCATCTCAAT 3889
QY 2641 ttgcccagcggaagaaatgggcaatcaatgctagcagcggttagaggaatccacccct 2700
D 3890 TTGCCGACGGGAAATGGGCAATCAATGCTACGAGCGGTAAGTAGAGAAATCCACCTT 3949
QY 2701 ggtcaagcagsgggaagtgltccaaagltacaggtatatactatgtagatccttcaagaag 2760
D 3950 GGTCAAGCAGAGGAGMACTGTCAAGTACACAGGTATATCTATGATGATCTTCATCAAGAG 4009
QY 2761 gtaagcccaagccacggtgaataaagaatagaataa 2794
D 4010 GTAGCCCAAGCACAGGTAAATAAGTAATAGAAA 4043

```

```

RESULT 13
US-08-474-140-10
: Sequence 10, Application US/08474140
: Patent No. 5721127
: GENERAL INFORMATION:
: APPLICANT: DEMER, PHILIPPE
: APPLICANT: AMORY, ANTOINE
: TITLE OF INVENTION: PULLULANSE, MICROORGANISMS WHICH
: TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESS: WILLIAM BRINKS HOFER GILSON & LIONE, P. C.
: STREET: 2000 K Street, N.W., Suite 200
: CITY: Washington

```

```

: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474.140
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Wilhem F. Gadiano, Esq.
: REGISTRATION NUMBER: 37,136
: REFERENCE/DOCKET NUMBER: 4121-41
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 429-0625
: TELEFAX: (202) 293-1850
: TELEX: 650 383-5605
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2784 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
US-08-474-140-10

```

```

Query Match 99.4%; Score 2776.4; DB 1; Length 2784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 gatgggaacagacacagcatctgtccactatlttcccgctggtgattacaact 60
D 1 GATGGGAACACGACACGATCTGTCCATTTTGGCTGTGTATATACACT 60
QY 61 tggagctctatgatttgcccaaaagagtgaggctgataactcaactcaaccg 120
D 61 TGGAGCTCTATGATGTGGCCAAAAGACGAGTGCGGCTGAATACGATTTCACACCG 120
QY 121 gctgactcttlttgagcgtgttgcaagtgctgatatccagaacccaagltcaggtaagg 180
D 121 GCTGACTCTTTTGGAGCTGTGCAAGTGTGATATTCAGGAACCCAAAGTCAGTAGGA 180
QY 181 attatcgttcgaccccaagattggacccaagaatgtgagcgctgacgcgtacaatgattca 240
D 181 ATTATCGTTTCGACCTCAAGATTGGACCAAGAATGTGAGCGCTACATAGATTTA 240
QY 241 agcaaaagaaatgagtggtggtctgtagaaggaacaagcacaatttttataatgaana 300
D 241 AGCAAAAGAAATGAGGTGTGGCTGTGAGAGGAACACGCCAAATTTTATATGAAAAA 300
QY 301 gatctgagatgacgtacaaacgcgtglaagcaacgcttatttgaatgcaaacag 360
D 301 GATCTGAGAGATGACGCTAAACCCGCTGAAGCAACGCTTATTTAGATCTTCAACACAG 360
QY 361 ggtcgtggttaaaacttagcagccgttaactcttgggggaaggnnaagcgcttaaggt 420
D 361 GTGCTGTTAAACTTAGCCAGCCGTTAACTCTTGGGAAGGNNNAAGCGGCTTTAGGGTT 420
QY 421 catgacacacagcaataaagataatccagtgacatcgtgaaagatcaagtcctgtgt 480
D 421 CATGACGACACGCAATTAAGATATTCACAGTACATCTGTGAAGATGCAAGCTTGTGT 480
QY 481 caaagatgtaacgcgtgttlttgcaaggtacccccaacataatlttgaaggttccagattg 540
D 481 CAAGATGTAAACGCTGTTTGGCAGGTACCTTCAACAATATTTTGGAGAGTCCGATGG 540
QY 541 ggaactgataatcaagtaacttataaaagggagacaaacaaactctcataacttctca 600
D 541 GCACCTGATTAATCACAGTACTTATTAATAAAGGTGACTTAACAAATCTATCAATTTCA 600

```



|||||  
Db 2761 GTAAGCCGACGACGCTAAAG 2784

## RESULT 14

US-08-477-630-10  
Sequence 10, Application US/08477630  
Patent No. 5721128  
GENERAL INFORMATION:  
APPLICANT: DEMEER, PHILIPPE  
APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULLULANSE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULLULANSE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,630  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-42  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2784 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-477-630-10

Query Match 99.48; Score 2776.4; DB 1; Length 2784;  
Best Local Similarity 100.00; Pred. No. 0;  
Matches 2783; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggagcaacagacatatttgcacatatttggccctgcggtgattatcaacct 60  
Db 1 GATGGGAACACGACACATATTTCATATTTTGGCTGCGTGATTAATCAACT 60  
QY 61 tggactatgatgtggcacaagaagcgggtgggctgtaataagattcaaccg 120  
Db 61 TGGACTATGATGTGGCCAAAAGACGAGGTGGGCTGATTAATCAATCAACCG 120  
QY 121 gctgactcttttggagctgttgcgaagtcgtataltccagaagaaccgaagtcagta 180  
Db 121 GCTGACTCTTTTGGAGCTGTGCAAGTCTGATATTCAGAAACCAAGTCAGTAGGA 180  
QY 181 attatcgttcgactcaagattggacacaagaatgtagcgctgacgcgtaactaata 240  
Db 181 ATTATCGTTGCGACTCAAGATTGGACCAAAAGTGTGAGCGCTGACCGCTACATAATTTTA 240  
QY 241 agcaaaagtaaatgagtgctgtgagctgtagaaggaacacgccaatttltatataa 300  
Db 241 AGCAAAAGTAATGAGTGCTGTGTAGAGGAACACGCCAATTTTATATATGAAAAA 300

QY 301 gatctgagatgacagctaaaccgcgtgaagcaacgcttattagatgcttcaaacag 360  
Db 301 GATGCTGAGGATGACAGCTAAACCCGCTGTAAAGCAACGGCTTATTGATGCTTCAAAACGAG 360  
QY 361 gtcgtggttaaaccttaagcagccgttactcttgggggaagggnaagggcttaccgtt 420  
Db 361 GTGCTGTTAAACCTTAAGCCAGCGGTTAACTCTTGGGGAAGNNNAAGCGGCTTACGGT 420  
QY 421 catgacgacacagcaaatagaataltccagtgacatcgtgaagagatgcaagcttctgt 480  
Db 421 CATGACGACACGCAAAATAGATATTTCAGTGACATCTGTAGAGATCAAGTCTTGCT 480  
QY 481 caaagatgaacccgctgttttggcaggtacccctccaacatalttggaggttcgattgg 540  
Db 481 CAAGATGTAAACCGCTGTTTGGCAGGTACCTTCAACATATTTTGGAGAGTTCCGATGG 540  
QY 541 gacctgttaatacagaacttatttaaaaaagggagacaaacatctatcaattctca 600  
Db 541 GCACCTGATAATCAGATGACTTATTAAAAAGGTCACTAACAAATCTTATCAATTCGA 600  
QY 601 ggaagatcttcctgaaaggaacataataaagtggtcttaaatgatatgctggaat 660  
Db 601 GGAGATCTTCCGAAAGAAACTACCAATTAAGTGGCTTTAAATGATAGCTGGAATAT 660  
QY 661 ccgagttaccacatcgcacacataaattaaacgtccctgcggtgcggtgcacagctact 720  
Db 661 CCGAGTTACCATCTGCACAACTTAATTAAACGTCCCTGCGGGGCTGCACACGCTACT 720  
QY 721 ttctgtaatactcgttccactcatgagctctatgacaataataatccctaatgcgat 780  
Db 721 TTTTCTGTAATTTCCGTCACCTCATGACGTATGACAAATTAATATCTTAATGGGAT 780  
QY 781 ttacaagtaagaagcgggttlaaaacggtatccgtgaagcttacccttaagggaaagctca 840  
Db 781 TTACAAGTGAAGCGGGTTAAACGATCTCGTACGCTTACTAGGGAAGATGCA 840  
QY 841 gatgtgagccatactctgtccatccaacagatggtatcaagcaagcaggtgatcct 900  
Db 841 GATGTAGCCATATCTGTCCATTCMAACAGATGGCTATACAGGCAAGCAGGTGATACCT 900  
QY 901 cgtaatgcttaattatcaacagatcactatccaaggtatgagatcttgggaatacct 960  
Db 901 CGTAAATGCTTAATTCAATCAGATGACTATTCAGGAGATGATTTGGCAATFACCTAT 960  
QY 961 acacagaagcaaacactttaaagctctggcaccacacatctcaacgtaaaatgctt 1020  
Db 961 ACACAGAAGCAAAACCTTAAAGTCTGGGACCAACTTCACTAATAATGTTCTT 1020  
QY 1021 cttatgacagtgcacaggggttcctgtaacaaaatctgtaacctatgacggcatcgccat 1080  
Db 1021 CTTTATGACAGTGCAAACGGGTTCTGTAAACAAAATGTAACCTATGACGGCATCGGCCAT 1080  
QY 1081 gctgtgtgggaagcaaggttlaacaaaaccttgaatgtglatatcgtatagggta 1140  
Db 1081 GGTGTGTGGGAAGCAACGGTTAATCAAAACCTTGAATGTGATTAATGAGTA 1140  
QY 1141 acagaagcaagctcctacccgaacggtctgttgaatcccttgaacactgcgttgcacaaat 1200  
Db 1141 ACAGGCAAGGCTCTACCGAAGCGGTGTGATCTTATGCAACTGCAATGCAACCAAT 1200  
QY 1201 ggaacgagagcagatgtgtgacacctgctaaacacagatccctgcgtgcgaacagtgat 1260  
Db 1201 GGAACGAGAGCATGTGTTGTGACCTGGCTTAAACAGATCCGTGCTGGCAACAGTGAT 1260  
QY 1261 aaacatatagcccaagaagatataagagtgaggtgtcatatagaatagatgtcgtagc 1320  
Db 1261 AAACATATTACCCCAAGAAATATAGAGATGAGGTATATTAATGAGATGTCCTGTGAC 1320  
QY 1321 ttctccatgacccctaaatccgggtatgaaataaaaggggaagctatttgccttcaagaa 1380  
Db 1321 TTTTCATTTGACCTTAATTCGGGTATGAATAAATAAGGAAGATTTGGCTCTTACAGAA 1380  
QY 1381 aaggaacaagggccctgacaaacgtlaaagacgggagtagatctccttaaaacaaactggg 1440



```

Db 1381 AAGGAAAGGAGGCGCTGACAGAGCTAAAGACGGGAGTAGATTCCTTAAACCACTGGG 1440
Qy 1441 attactactgltcagactatgagcgttcttcgcatctaacagtgctgcatgaaactatca 1500
Db 1441 ATTACTATGTTGACGCTTATGCGCTGTTTGGCATCTAAACAGTGTCGATGAAACGATCCA 1500
Qy 1501 acccaagaataatctgggttatgagccctcgcaactatgagtctcctgaaggagatgct 1560
Db 1501 ACCCAAGATATATGGGGTATTGACCTCGCAACATGATGATCTCTGAAAGGCGCATATGCT 1560
Qy 1561 acbaatgcgaatgttaatgctcgtatbaaagagtttaaggaaatggtctcttcctcat 1620
Db 1561 ACAATGCGAATGCTAATGCTCGATATAAGAGATTAAAGCAATGATCTTTTCACTCCAT 1620
Qy 1621 cgltaacacaltgaggttcaacalbgaltgctataatcatatcattctgcaacgaatc 1680
Db 1621 CGTCAACACATGCGGCTTAAACATGATGTTGCTATATCATACCTTTGCCACGCCAATC 1680
Qy 1681 tctgactctgataaataatctgacagaatatactacgtacagatgacagtaattata 1740
Db 1681 TCTGACTTCTGATTAATATGTAACCAATATTATACCGTACGATGATGACGATATATATA 1740
Qy 1741 ccaacgagatcaggtactggaatgaatctgacagcgaagcgaatggtctcaaaattc 1800
Db 1741 CCAACGATCAGGCTACTGCAATCAATTCGACNCAAGGCCAATGCTTCAAAATTT 1800
Qy 1801 attatgattcccttaagatctgggtcaatgagatcatatctgaggtctccgttctgac 1860
Db 1801 ATTATGATTCCCTTAACTGATGTTGGCTCAATGATCATATTGACGCGCTTCCGTTTGGAC 1860
Qy 1861 ttatgagcgtctctgtaaaagacagatgctccaaagctcctctggaggtctcatgcatc 1920
Db 1861 TTATGAGCGTCTGTTGAAAAAGACAGATGTCCAAAGCTCCTCTGGAGCTTCAATGCTATT 1920
Qy 1921 aatcagaagattgcaacttcaagtgagtgagcagtgagcaggtctgcaactgacagat 1980
Db 1921 AATCCAGAAATTCGACTTTACGTCGACCATGACGCGGTGCAACCTCTGCACTGCCAGAT 1980
Qy 1981 gatcagctctgacaaaagagagctcaaaaagcaltggagtagcgltcttaagaaat 2040
Db 1981 GATCAGCTCTGACAAAAGAGCTCAAAAAGCATGCGAGTACGCGTCTTAAATGACAAAT 2040
Qy 2041 ttgcaaaagcgttgagcagcaatgcttctgactctcgcctcaaggtcttctgagcaggt 2100
Db 2041 TTGCAAAAGCGTGGAGCGCAATGCTTGTGATCTTCCCTCAAGGTTTTCGACAGGT 2100
Qy 2101 gcaacagcttaactgaatgaatgaatgagcgttgagggaggtatgaatgacttacc 2160
Db 2101 GCAACAGCTTAATGATGCAATTAAGATGCGCTTGAGGCGAATTAATGACTTTACC 2160
Qy 2161 tcttaccaggttgagcaatlaactatgctacaaagtgatgaactacaccccttggagc 2220
Db 2161 TCTTACACAGCTGAGCAATTAATGATGTCAGAACTCATATACTACACCTTTGGAGC 2220
Qy 2221 aaaaatgacctgaagcaatgcttaatgactcgaagcgaatgaggttaaatgagtgaatc 2280
Db 2221 AAAATAGCCCTTAGCAATTCCTAATGATTCGAAAGCGGATGGATTAATAAGATGAATC 2280
Qy 2281 gcaacagcagctgctatgactcgaagcgttctcattctgcaagcggggaagaatg 2340
Db 2281 GCAACAGCAGCTTGTATGACTCTGCAAGGCGTTCATTCATGCAAGCGGCGGAAGAATG 2340
Qy 2341 ctctgtanaaaagcggcgaacagcaatagtlataatgcaagcgatgctgcatagttt 2400
Db 2341 CTTCGTANAAAAGCGCCGACAGCAATAGTTATATCAGCGCATCGGTCATAGCTTT 2400
Qy 2401 gatttgagcaggaagatcgaatcgaatgcttctcaactatctagcgggtcaatccac 2460
Db 2401 GATTGAGCGAGAAAGCTCAATATCCAGATGTTTCAATATTAATACGGGCTAATCCAC 2460
Qy 2461 ctctgctctatcaaccagcctctccgcatgacgaacgaatgaatcgaatagcaactc 2520

```

```

Db 2461 CTTCGCTTGTATGATCCAGCCATCCGATGACGACAGCTAATGAATCAATAGCCACTC 2520
Qy 2521 caattcctaaatagtcagagaaacagtgagccatgtaattaaactgactgtaataaa 2580
Db 2521 CAATTCTTAATAGTCCAGAGAACACAGTGGCGATATGATTAATCAATGATTAATAA 2580
Qy 2581 gacaaatggggaatatacatgttcttataaaccbaaataaactgtagcaacatcaat 2640
Db 2581 GACAAATGGGGAAATATCATGTGTTTATTAACCAATATAAATCTATGACCAATCAAT 2640
Qy 2641 ttgcccagcgggaatgaggaatcaatgctcagagcggtaaggaagaatccacctt 2700
Db 2641 TTGCGGAGCGGGAATGCGCAATCAATGCTAAGCAGCGGTAAAGTAAGATCCACCTT 2700
Qy 2701 ggtcaagcagaggaagtgctccaaagtaatacatatcatgataatgactcctcaagaag 2760
Db 2701 GGTCAACGACAGGGAATGTGTCAGAGTACAGAGTATATCTATGATGATCTTCAATCAAGAG 2760
Qy 2761 gtaagccagacagcaggttaaaaag 2784
Db 2761 GTAAGCCGACAGCACGCGTAAAGAG 2784

RESULT 15
US-08-472-293-10
: Sequence 10, Application US/08472293
: Patent No. 5731174
: GENERAL INFORMATION:
: APPLICANT: DEMER, PHILIPPE
: APPLICANT: AMORY, ANTOINE
: TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH
: TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
: STREET: 2000 K Street, N.W., Suite 200
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/472,293
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Wilhelm F. Gagliano, Esq.
: REGISTRATION NUMBER: 37,136
: REFERENCE/DOCKET NUMBER: 4121-44
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 429-0625
: TELEFAX: (202) 293-1850
: TELEX: 650 383-5605
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2784 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: US-08-472-293-10

Query Match 99.4%; Score 2776.4; DB 1; Length 2784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gatgggaacagcaacagatcatgttccactatttttgcctctgctgctgattatcaact 60

```



|||||  
Db 1 GATGGGAACAGACACATCATTTGTCACATTTTTCCTGCGTGATTAACCT 60  
Oy 61 tggagctatgatctggtgccaagaagcggtggtgctgataatcattcaacacg 120  
|||||  
Db 61 TGGAGTCTATGATGATGCGCCAAAAGACGAGGTGGGCTGTAATGACATTTCAATACCG 120  
Oy 121 gctgactcttltggagctgtltgcaaglyctgataltccaagaaacccaagtcagta 180  
121 gctgactcttltggagctgtltgcaaglyctgataltccaagaaacccaagtcagta 180  
Db 121 GCTGACTCTTTTGGAGCTGTTCGAAGTGTGATTTCCAGGAAACCAAGTCAGTAGGA 180  
Oy 181 attatcgctgcactcaaatgtgagccaaagaatgtgagcgctgagccgtacataa 240  
181 attatcgctgcactcaaatgtgagccaaagaatgtgagcgctgagccgtacataa 240  
Db 181 ATTATCGTTCGACTCAAGATTGGACCAAGATGTGACGCGTGACCGCTTACATTA 240  
Oy 241 agcaagaagaatgaggtgtggtctgtagaagaagaacccaattttataatga 300  
241 agcaagaagaatgaggtgtggtctgtagaagaagaacccaattttataatga 300  
Db 241 AGCAAGAAGAATGAGGTGTGCTTGTGCTTACAGCAAGCAACCAATTTTATATGAAAA 300  
Oy 301 gctgctgagatgcaagctaaacccgctgtlaagcaagctlatltagatgctcaacag 360  
301 gctgctgagatgcaagctaaacccgctgtlaagcaagctlatltagatgctcaacag 360  
Db 301 GATGCTGAGATGCAAGTAAACCGCTGTAAAGCAAGCTTATTTAGATGCTCAAAACCG 360  
Oy 361 gltgctglttaacttaagccagcgtlaactccttggggaagannaagcgcttca 420  
361 gltgctglttaacttaagccagcgtlaactccttggggaagannaagcgcttca 420  
Db 361 GTGCTGTTAAACTTACGACGCGCTTAACTCTTGGGGAAGNNNAAGCGCTTTACG 420  
Oy 421 catgagcagaagcaaaagaatatactcaatgacatcgttgaagaagcaagttt 480  
421 catgagcagaagcaaaagaatatactcaatgacatcgttgaagaagcaagttt 480  
Db 421 CATGACGACACAGCAAAATAGATATTCAGTACATCTGTAAAGATGCCAAGTCTGT 480  
Oy 481 caagatgtaaacgcgtgtlttgcaagtaacctccaacatatttggaggtlccag 540  
481 caagatgtaaacgcgtgtlttgcaagtaacctccaacatatttggaggtlccag 540  
Db 481 CAAGATGTAAACCGCTGTGCTTGGCAGTAGTACCTTCAACATATTTTGGAGTTC 540  
Oy 541 gcaactgtaatacagtaacttataaagaaglygaactaactctataacttca 600  
541 gcaactgtaatacagtaacttataaagaaglygaactaactctataacttca 600  
Db 541 GCACCTGTATATCAACAGTACTTATTAATAAAGTGACTAACATCTATCAATTC 600  
Oy 601 ggaagcttccctggaagaagaactacaataaagtgtgcttaagtgtgagtg 660  
601 ggaagcttccctggaagaagaactacaataaagtgtgcttaagtgtgagtg 660  
Db 601 GGAGATCTTCCGTAAGGAACCTACCAATATTAAGTGCTTAAATGATAGCTGAAT 660  
Oy 661 ccgagttaccactcgtgaacaacataatlttaacagtcctcgtccggtgcaac 720  
661 ccgagttaccactcgtgaacaacataatlttaacagtcctcgtccggtgcaac 720  
Db 661 CGGATTTACCATCTGACAACATTAATTAACATGCTCCGCGGGGTGCACACG 720  
Oy 721 ttttgcgtatattcgtlccaactcaatgtcgtatgacacaaatlaaataat 780  
721 ttttgcgtatattcgtlccaactcaatgtcgtatgacacaaatlaaataat 780  
Db 721 TTTTGTATATCCCTGACACTCATGACGTATGACACAAATTAATCTATATG 780  
Oy 781 ttacaagtagaagaagcggtttaaagcgtatcgtgagcggttaacttaaga 840  
781 ttacaagtagaagaagcggtttaaagcgtatcgtgagcggttaacttaaga 840  
Db 781 TTACAAGTAGAAGCGGGGTAAACCGATCTGCTGACGCTTACTAGGGAAG 840  
Oy 841 gatgtgaagccatctcgtlccaactcaacagaatgtctalcaaggcaagcag 900  
841 gatgtgaagccatctcgtlccaactcaacagaatgtctalcaaggcaagcag 900  
Db 841 GATGTGAGCGCATCTGTCATCAATCAAGATGGCTATCAGCAAGAGGTGATAC 900  
Oy 901 cgtatgtgcttaattcatcatcagtaactatltcaagaagatgacttvgga 960  
901 cgtatgtgcttaattcatcatcagtaactatltcaagaagatgacttvgga 960  
Db 901 CGTATGTGCTTAATTCATACAGTACTATTCAGGAGATGATCTTGGGAAT 960  
Oy 961 acacagaagaagaacacactttaaagtltgggaacaaacttaacaaatga 1020  
961 acacagaagaagaacacactttaaagtltgggaacaaacttaacaaatga 1020  
Db 961 ACACAGAAGAAGCAACACTTTAAAGTCTGGGCAACAACTTCTACTCMAAT 1020  
Oy 1021 ctltatgacagtgcagcggtgtcgtlaacaaagaatcgtlaactaagcagc 1080  
1021 ctltatgacagtgcagcggtgtcgtlaacaaagaatcgtlaactaagcagc 1080  
Db 1021 CTTTATGACAGTGCAGCGGGTCTGTAAACAAATCGTACGTATGACGCGCAT 1080  
Oy 1081 ggtgtgtgggaagcaaggttcaataacacacttgaagaatgtgatatcag 1140  
1081 ggtgtgtgggaagcaaggttcaataacacacttgaagaatgtgatatcag 1140

|||||  
Db 1081 GGTGTGGAAGCAACGGTTATATCAAAACCTTGAAAATTTGATATTAC 1140  
Oy 1141 acagccaaagcgtctaacccgaagcgtgtgtgatacttaltgaacatgca 1200  
1141 acagccaaagcgtctaacccgaagcgtgtgtgatacttaltgaacatgca 1200  
Db 1141 AAGGCCAAAGCGCTTACCGGAACGGCTGTGATTCCTTATGCAACGTCCACCAAT 1200  
Oy 1201 ggaacgaagcagatgtgtgtgacgtgtgtaaaacagatcctgtgtgtg 1260  
1201 ggaacgaagcagatgtgtgtgacgtgtgtaaaacagatcctgtgtgtg 1260  
Db 1201 GGAACGAGCGCATGTTGTGACCTGTGCTAAACAGATCGCTGGCGGAAC 1260  
Oy 1261 aaacattagccaaagaatataagaagaatgaggtgcatatgaagaatga 1320  
1261 aaacattagccaaagaatataagaagaatgaggtgcatatgaagaatga 1320  
Db 1261 AAACATATTACCCCAAGAAATATAGAGATGAGTATCTTATGAATGATG 1320  
Oy 1321 tttccatgaccccaattcgtgtgtatgaagaataaagggaatatttgc 1380  
1321 tttccatgaccccaattcgtgtgtatgaagaataaagggaatatttgc 1380  
Db 1321 TTTTCATTGACCCCTAATTCGGGTATGAAAAATTAAGGAGATTTGGCT 1380  
Oy 1381 aaaggaacaaaggccctgacaacglaaagacggtgatalcctlaaaca 1440  
1381 aaaggaacaaaggccctgacaacglaaagacggtgatalcctlaaaca 1440  
Db 1381 AAAGGAACAAAGGCCCTGACAAAGTAAGACGGGATGATTTCTTAAAC 1440  
Oy 1441 attactcaatgctcagcttatgctgttltcgcatccaagatgtgataact 1500  
1441 attactcaatgctcagcttatgctgttltcgcatccaagatgtgataact 1500  
Db 1441 ATTACTCATGTTTCAAGCTTATGCTGTGTTTCGATCTAACAGTGTGAT 1500  
Oy 1501 acccaagaatacttgggttatgacccctgcacactagatgttccgtaagg 1560  
1501 acccaagaatacttgggttatgacccctgcacactagatgttccgtaagg 1560  
Db 1501 ACCCAAGATTAATTTGGGTTATGACCTCCCAACTATGATGTTCTTAAG 1560  
Oy 1561 acnaatgcaatgltatgctgtatlaaagaagttlaaggaatgttcttca 1620  
1561 acnaatgcaatgltatgctgtatlaaagaagttlaaggaatgttcttca 1620  
Db 1561 ACAAAATGCCAATGTATATGCTGTATTAAGAAAGTAAAGAAATGTTCT 1620  
Oy 1621 cgtgaacacatgtgggttataatgtgtgtgtcctataactaactcct 1680  
1621 cgtgaacacatgtgggttataatgtgtgtgtcctataactaactcct 1680  
Db 1621 CGTGAACACATTTGGGGTTAATGATGATGTTGTATATCAATCTTCC 1680  
Oy 1681 tctgactcgtataaatacttaccagaatataatcagtaactgaactga 1740  
1681 tctgactcgtataaatacttaccagaatataatcagtaactgaactga 1740  
Db 1681 TCTGACTTGTATTAATTTGACCAATATTAATACCGTATGACAGTAAT 1740  
Oy 1741 ccaacgtaacagtaactgtgaataatgtgacangcmaagcgaatgttca 1800  
1741 ccaacgtaacagtaactgtgaataatgtgacangcmaagcgaatgttca 1800  
Db 1741 CCAACGATACAGTACTGGAATAATGCAATGCAACGCAATGTTCAAAAT 1800  
Oy 1801 attatgtatcccttaagatattgtgtcgaatgatacatatgaagcctc 1860  
1801 attatgtatcccttaagatattgtgtcgaatgatacatatgaagcctc 1860  
Db 1801 ATTATGATTTCCCTTAAAGTATGTTGGTCAATGATATCATATTTGAC 1860  
Oy 1861 ttaatgtgcgtgtcttgaagaagaacagatgtccaaagcgtcctgagat 1920  
1861 ttaatgtgcgtgtcttgaagaagaacagatgtccaaagcgtcctgagat 1920  
Db 1861 TTATGTGCGCTGTGTAAGAAAGACAGATGTCCAAAGCTGCGAGCTTCA 1920  
Oy 1921 aatcagaagaatgtgacttlaaggtgaagcagatgtgaacgtcgtgacat 1980  
1921 aatcagaagaatgtgacttlaaggtgaagcagatgtgaacgtcgtgacat 1980  
Db 1921 AATCCAGGAATTCACATTTACGGTGAAGCATGACGCGGTGCAACCT 1980  
Oy 1981 gatcagcttctgacaaagaagcgtcaaaagaatgtgggttagcgtgtt 2040  
1981 gatcagcttctgacaaagaagcgtcaaaagaatgtgggttagcgtgtt 2040  
Db 1981 GATCAGCTTCTGCAAAAGAGCTCAAAAAGGATGGAGTACGGTGTATTA 2040  
Oy 2041 ttacgaagaacgtgtgagcgaatgtccttgcattcctcgaagtttgcag 2100  
2041 ttacgaagaacgtgtgagcgaatgtccttgcattcctcgaagtttgcag 2100  
Db 2041 TTACGAAGACCGCTTGAACCGCAATGCTTTGATCTTCCGCTTAAGTT 2100  
Oy 2101 gcaacagcttlaactgaatgaatlaagaaatgtgtgagggagatlaatga 2160  
2101 gcaacagcttlaactgaatgaatlaagaaatgtgtgagggagatlaatga 2160  
Db 2101 GCAACAGCGCTTAACTGATCAATTAAGAATGGCGTGAAGGAGTATTA 2160  
Oy 2161 tcttcaacaggtgtgagcaatlaactatgtcacaagaatcgtataactac 2220  
2161 tcttcaacaggtgtgagcaatlaactatgtcacaagaatcgtataactac 2220  
Db 2161 TCTTACCAAGGTGAGCAATTAATATGTATCAAAAGTATGATTAACCC 2220

QY 2221 aaaaagccctaaagcaatccatgattccgaagcgatcgattaaatgatgaactc 2280  
 |||||||  
 Db 2221 AAAATAGCCCTTAACCAATCCTAATGATTCGAAACGGATCGATTAAATGATGAACCTC 2280  
 QY 2281 gcacaaagcagltgltatgaccctcacaagcgltccatcctgcaagcgagggaagaatg 2340  
 |||||||  
 Db 2281 GCACAAAGCAGTGTGTAAGACCTCACAAGCGCTTCATTGATGCAAGCGGGGAAGAAATG 2340  
 QY 2341 ctctgtanaaaaggcgcaacgacaatagtataatcgaagcgatgcggtcaatgatt 2400  
 |||||||  
 Db 2341 CTTCGTANAAAGGCGCAACGACAAATAGTATATGACAGCGATGCGTCAATGAGTTT 2400  
 QY 2401 gatggagcgagaaagctcaatcaccagatgtttcaactatlaagggctaatccac 2460  
 |||||||  
 Db 2401 GATTGACGACGAAAGCTCAATATCCAGATCTTTCAACTATTTATAGCGGCTAATCCAC 2460  
 QY 2461 ctctgltcctgacccagcctccgcatgagacagctaaatgaatcctatgaccactc 2520  
 |||||||  
 Db 2461 CTTCGTCTTGATCACCAGCCTTCGCATGACGACGATATGAATCAATAGCCACCTC 2520  
 QY 2521 caattcctaaatagtcacagagaacacagtgagcctaagaatbaactgaatcgttaataa 2580  
 |||||||  
 Db 2521 CAATTCTTAATAGTCCAGAGAACACAGTGCGCTATGATTAATTAATCAATGATTAATAAA 2580  
 QY 2581 gacaaatggggaatlaatcattgltgttlaaaccaataaaactgtagcaacatcaat 2640  
 |||||||  
 Db 2581 GACAAATGGGGAATATCATTTGTTTATTAACCAATAAATCTGTAGCAACCATCAAT 2640  
 QY 2641 ttgcgagcgaggaaatgggaatcaatgctlaagcggtlaagtaggaagaatccacct 2700  
 |||||||  
 Db 2641 TTGCGAGCGGGAATGCGCAATGCTAGCGAGCGTAAGTAGGAGAAATCCACCCTT 2700  
 QY 2701 ggtcaagcagaggaagtgtccaaagtacaggtatatactatgataatccctcaagaag 2760  
 |||||||  
 Db 2701 GGTCAAGCAGAGGAGAGTGTCCAAGTACAGGTATATCTATGATGATCCTTCATCAAGAG 2760  
 QY 2761 gtaagccagacacacggtlaaaag 2784  
 |||||||  
 Db 2761 GTAAGCCACACACGCGTAAAAAG 2784

Search completed: April 22, 2001, 07:58:35  
 Job time: 5669 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2001, 06:52:19 ; Search time 44.44 Seconds  
(without alignments)  
8821.730 Million cell updates/sec

Title: US-09-262-126C-1

Perfect score: 2794  
Sequence: 1 gatgggaacacgacacacgat.....cgttaaaagtaataagaaaa 2794

Scoring table: IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 48221 seqs, 70157065 residues

Total number of hits satisfying chosen parameters: 96442

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_MA\_New:\*  
1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2788	99.8	2794	5	US-09-262-126C-1
2	52	1.9	2706	5	US-09-463-238-2
3	45.2	1.6	2806	5	US-09-463-238-11
4	39.4	1.4	2843	5	US-09-463-238-1
5	39.4	1.4	2843	5	US-09-463-238-10
6	37.4	1.3	2634	5	US-09-463-238-3
7	35.8	1.3	264546	6	US-60-248-505-202
8	35	1.3	91051	6	US-60-248-505-164
9	35	1.3	133937	6	US-60-248-822-8
10	34.4	1.2	2794	5	US-09-262-126C-1
11	34.2	1.2	22788	6	US-60-248-505-127
12	34.2	1.2	30167	6	US-60-248-505-611
13	34.2	1.2	55248	6	US-60-248-505-254
14	34.2	1.2	61635	6	US-60-248-505-244
15	34.2	1.2	150471	6	US-60-248-505-148
16	33.6	1.2	1389	5	US-09-463-238-12
17	33.4	1.2	35945	6	US-60-248-505-547
18	33.4	1.2	49055	6	US-60-248-505-77
19	33.4	1.2	96187	6	US-60-248-505-52
20	33.4	1.2	506113	5	US-09-739-449-218
21	33.2	1.2	500	4	US-08-276-163D-13580
22	33	1.2	36506	6	US-60-248-505-107
23	32.8	1.2	397658	5	US-09-813-320-3
24	32.6	1.2	38593	6	US-60-248-505-425
25	32.6	1.2	75545	6	US-60-248-505-548
26	32.6	1.2	96768	6	US-60-248-505-516
27	32.6	1.2	111782	6	US-60-248-505-62

c	28	32.6	1.2	141671	6	US-60-248-505-256	Sequence 256, App
c	29	32.2	1.2	82938	5	US-09-818-657-3	Sequence 3, Appl1
c	30	32.2	1.2	147644	6	US-60-248-505-6	Sequence 6, Appl1
c	31	32	1.1	26076	6	US-60-248-505-575	Sequence 575, App
c	32	32	1.1	44081	6	US-60-248-505-282	Sequence 282, App
c	33	31.8	1.1	33351	5	US-09-543-679A-2874	Sequence 2874, App
c	34	31.8	1.1	40298	5	US-09-543-679A-2878	Sequence 2878, App
c	35	31.4	1.1	264107	6	US-60-248-505-207	Sequence 207, App
c	36	31.4	1.1	35167	6	US-60-248-505-78	Sequence 78, Appl
c	37	31.2	1.1	24273	6	US-60-248-505-24	Sequence 24, Appl
c	38	31.2	1.1	118711	6	US-60-248-505-389	Sequence 389, App
c	39	31.2	1.1	240077	6	US-60-248-505-7	Sequence 7, Appl1
c	40	31	1.1	39467	6	US-60-248-505-214	Sequence 214, App
c	41	31	1.1	45333	6	US-60-248-505-604	Sequence 604, App
c	42	31	1.1	449171	6	US-60-248-505-42	Sequence 42, Appl
c	43	30.8	1.1	4324	6	US-60-248-505-380	Sequence 380, App
c	44	30.8	1.1	68264	6	US-60-248-505-506	Sequence 506, App
c	45	30.8	1.1	120994	6	US-60-248-505-236	Sequence 236, App

## ALIGNMENTS

RESULT 1  
US-09-262-126C-1  
; Sequence 1, Application US/09262126C  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Brian S.  
; APPLICANT: Shetty, Jayarama K.  
; TITLE OF INVENTION: Modified Forms of Pullulanase  
; FILE REFERENCE: GC396-2  
; CURRENT APPLICATION NUMBER: US/09/262,126C  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2794  
; TYPE: DNA  
; ORGANISM: Bacillus deramificans  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)...(2794)  
; OTHER INFORMATION: n = A, T, C, or G  
US-09-262-126C-1

QY	1	gatgggaacacgacacacgatcattgtccactattttgcccctgtgtattacaact	60
DB	1	gatgggaacacgacacacgatcattgtccactattttgcccctgtgtattacaact	60
QY	61	tggaattctatgtgtgtgcacaaagacggaaggttggggctgtaatacattcaacacg	120
DB	61	tggaattctatgtgtgtgcacaaagacggaaggttggggctgtaatacattcaacacg	120
QY	121	gctgactcttttggagctgttcaagtgctgatatccaggaacccaagtcagttaga	180
DB	121	gctgactcttttggagctgttcaagtgctgatatccaggaacccaagtcagttaga	180
QY	181	attatcgttcgacacccaatttggaacaaagatgtgagcgtgacccgtacataagtta	240
DB	181	attatcgttcgacacccaatttggaacaaagatgtgagcgtgacccgtacataagtta	240
QY	241	agcaaaaggaatgaggtgtgctgtgtagaaggaacagcaaatctttataatgaaaa	300
DB	241	agcaaaaggaatgaggtgtgctgtgtagaaggaacagcaaatctttataatgaaaa	300
QY	301	gatgtcgaagatgacgcgttaaacccgtgtgaagcaagctatttgatgtctcaacacg	360
DB	301	gatgtcgaagatgacgcgttaaacccgtgtgaagcaagctatttgatgtctcaacacg	360

QY 361 gtctgtgttaaaccttagccagccgtttaactctgtgggaagmnaagcgttaagctc 420  
|||||  
Db 361 gtctgtgttaaaccttagccagccgtttaactctgtgggaagmnaagcgttaagctc 420  
QY 421 catgaacagacagagaaataagagatattccagtgagacatctgtgaagagtgcaagcttgc 480  
|||||  
Db 421 catgaacagacagagaaataagagatattccagtgagacatctgtgaagagtgcaagcttgc 480  
QY 481 caagagtgtaaccgcgtctgtctgtgcagagcttccacaatatcttttggaagcttccgattgg 540  
481 caagagtgtaaccgcgtctgtctgtgcagagcttccacaatatcttttggaagcttccgattgg 540  
QY 541 gccccttaataatcacagttactttaataaagaagtgagctaacatctctcttaatttca 600  
|||||  
Db 541 gccccttaataatcacagttactttaataaagaagtgagctaacatctctcttaatttca 600  
QY 541 gcaaccgatatatacagtaactttaataaagaagtgagctaacatctctcttaatttca 600  
|||||  
QY 601 ggaagatctctcgtgaagagaaatacacaataaagtgagcttaaaatgaatagctgaaat 660  
|||||  
Db 601 ggaagatctctcgtgaagagaaatacacaataaagtgagcttaaaatgaatagctgaaat 660  
QY 661 ccgaagtatacccatctgagacaataatlaataacagttccctgcggcgggtgcacacgttca 720  
|||||  
Db 661 ccgaagtatacccatctgagacaataatlaataacagttccctgcggcgggtgcacacgttca 720  
QY 721 ttcttcgtaatttcctgtccactcatcgagcttatagacacaataataatcttaattgcgagat 780  
|||||  
Db 721 ttcttcgtaatttcctgtccactcatcgagcttatagacacaataataatcttaattgcgagat 780  
QY 781 ttaacaaatgaagaagcgggtgttaaaacagatctcgtagagcttaactcaggggaagaaacca 840  
|||||  
Db 781 ttaacaaatgaagaagcgggtgttaaaacagatctcgtagagcttaactcaggggaagaaacca 840  
QY 841 gatgtgagacatactctgtccatccaagaagatgagctatcaagcaaaagcagtgatacct 900  
841 gatgtgagacatactctgtccatccaagaagatgagctatcaagcaaaagcagtgatacct 900  
Db 841 gatgtgagacatactctgtccatccaagaagatgagctatcaagcaaaagcagtgatacct 900  
QY 901 cgttaattgtcttaattatcaacagttactatactcagagagatgactctgggaataacctat 960  
|||||  
Db 901 cgttaattgtcttaattatcaacagttactatactcagagagatgactctgggaataacctat 960  
QY 961 acacaggaagaaacaaacaccccttaaaagtcgtggcacaaacctctcaatgaaatgtctctc 1020  
|||||  
Db 961 acacaggaagaaacaaacaccccttaaaagtcgtggcacaaacctctcaatgaaatgtctctc 1020  
QY 1021 ctcttaagacagtgacacaggtctctgtacaaaatcgttacctatgacgacatcgagccat 1080  
|||||  
Db 1021 ctcttaagacagtgacacaggtctctgtacaaaatcgttacctatgacgacatcgagccat 1080  
QY 1081 ggtctgtgtggaagcaacggtttaatacaaaccttgaaatactgtaatacagttatgagata 1140  
|||||  
Db 1081 ggtctgtgtggaagcaacggtttaatacaaaccttgaaatactgtaatacagttatgagata 1140  
QY 1141 acagagccaaagctcttaaccggaacggtctgtatgactctatgcaactggaatctgacaaat 1200  
|||||  
Db 1141 acagagccaaagctcttaaccggaacggtctgtatgactctatgcaactggaatctgacaaat 1200  
QY 1201 ggaagagagagagcagatgtgtgtgagacgtgacataaaacagatccctgcgtggaacagtgat 1260  
1201 ggaagagagagagcagatgtgtgtgagacgtgacataaaacagatccctgcgtggaacagtgat 1260  
Db 1201 ggaagagagagagcagatgtgtgtgagacgtgacataaaacagatccctgcgtggaacagtgat 1260  
QY 1261 aaacatacttcgcgaagaagaaatacagaagatgaggttcaatctgaataatgagtgacggagac 1320  
|||||  
Db 1261 aaacatacttcgcgaagaagaaatacagaagatgaggttcaatctgaataatgagtgacggagac 1320  
QY 1321 ttcttcacattgaaaccttaattctcgtgtatgaaataataaagggagagatctgtctcttaagaa 1380  
1321 ttcttcacattgaaaccttaattctcgtgtatgaaataataaagggagagatctgtctcttaagaa 1380  
Db 1321 ttcttcacattgaaaccttaattctcgtgtatgaaataataaagggagagatctgtctcttaagaa 1380  
QY 1381 aaagagaaacaaagggcccttgacaaagctaaagaacggggagatagatctccttaaaacacttg 1440  
1381 aaagagaaacaaagggcccttgacaaagctaaagaacggggagatagatctccttaaaacacttg 1440  
Db 1381 aaagagaaacaaagggcccttgacaaagctaaagaacggggagatagatctccttaaaacacttg 1440

QY 1441 attactaatgttcagagcttaagccgtgttctgcatactaaacagtgatgcataacatgcatacca 1500  
1441 attactaatgttcagagcttaagccgtgttctgcatactaaacagtgatgcataacatgcatacca 1500  
Db 1441 attactaatgttcagagcttaagccgtgttctgcatactaaacagtgatgcataacatgcatacca 1500  
QY 1501 acccaagataatltgggtltatgaaacctcgcgaactagatgtctctgaagggcaatagatc 1560  
1501 acccaagataatltgggtltatgaaacctcgcgaactagatgtctctgaagggcaatagatc 1560  
Db 1501 acccaagataatltgggtltatgaaacctcgcgaactagatgtctctgaagggcaatagatc 1560  
QY 1561 acaaatgacgaatgtgaaatgtctgtatataaagaggtttaaggaatagttcttctccatcac 1620  
1561 acaaatgacgaatgtgaaatgtctgtatataaagaggtttaaggaatagttcttctccatcac 1620  
Db 1561 acaaatgacgaatgtgaaatgtctgtatataaagaggtttaaggaatagttcttctccatcac 1620  
QY 1621 cgtgaacaacattgggtgttaacatgagatgtgtctcaataataacatcttgcacagcaatc 1680  
1621 cgtgaacaacattgggtgttaacatgagatgtgtctcaataataacatcttgcacagcaatc 1680  
Db 1621 cgtgaacaacattgggtgttaacatgagatgtgtctcaataataacatcttgcacagcaatc 1680  
QY 1681 tctgaacttcgataaaatctgataccagaataatattatataccgttccgatatccaggttaataata 1740  
1681 tctgaacttcgataaaatctgataccagaataatattatataccgttccgatatccaggttaataata 1740  
Db 1681 tctgaacttcgataaaatctgataccagaataatattatataccgttccgatatccaggttaataata 1740  
QY 1741 ccaacggatcaggttactggaatgaaatgtgcangcngaaagccaaatgttccaaataatc 1800  
1741 ccaacggatcaggttactggaatgaaatgtgcangcngaaagccaaatgttccaaataatc 1800  
Db 1741 ccaacggatcaggttactggaatgaaatgtgcangcngaaagccaaatgttccaaataatc 1800  
QY 1801 attactgaatcccttaagatctgggtccaatgagatcataatctgaaggtctccgttttgac 1860  
1801 attactgaatcccttaagatctgggtccaatgagatcataatctgaaggtctccgttttgac 1860  
Db 1801 attactgaatcccttaagatctgggtccaatgagatcataatctgaaggtctccgttttgac 1860  
QY 1861 ttaatgagcgtcgtcttggaagaaagacagatgtccaaagctgcctcgtgagcttccatgacatc 1920  
1861 ttaatgagcgtcgtcttggaagaaagacagatgtccaaagctgcctcgtgagcttccatgacatc 1920  
Db 1861 ttaatgagcgtcgtcttggaagaaagacagatgtccaaagctgcctcgtgagcttccatgacatc 1920  
QY 1921 aatccaggaattgacatcttaacgttgagccaatgagcgggtggaacctctgcacttgacagat 1980  
1921 aatccaggaattgacatcttaacgttgagccaatgagcgggtggaacctctgcacttgacagat 1980  
Db 1921 aatccaggaattgacatcttaacgttgagccaatgagcgggtggaacctctgcacttgacagat 1980  
QY 1981 gatcagcttctgcgaagaagaagagctccaaagaagcagtggagatagcgtgttaatgacaat 2040  
1981 gatcagcttctgcgaagaagaagagctccaaagaagcagtggagatagcgtgttaatgacaat 2040  
Db 1981 gatcagcttctgcgaagaagaagagctccaaagaagcagtggagatagcgtgttaatgacaat 2040  
QY 2041 ttaacgaagacggtltgagacagcaatgtcttgaattctccgcctcgaaggttlttgacagaggt 2100  
2041 ttaacgaagacggtltgagacagcaatgtcttgaattctccgcctcgaaggttlttgacagaggt 2100  
Db 2041 ttaacgaagacggtltgagacagcaatgtcttgaattctccgcctcgaaggttlttgacagaggt 2100  
QY 2101 gcaacaggtcttaactgtagcaatlaagaatgagcgtltgaaggaggatlaatgacttacc 2160  
2101 gcaacaggtcttaactgtagcaatlaagaatgagcgtltgaaggaggatlaatgacttacc 2160  
Db 2101 gcaacaggtcttaactgtagcaatlaagaatgagcgtltgaaggaggatlaatgacttacc 2160  
QY 2161 tcttcaacaggtgtagacaataatgaatgatacagaagatgataatacctacaccccttgagac 2220  
2161 tcttcaacaggtgtagacaataatgaatgatacagaagatgataatacctacaccccttgagac 2220  
Db 2161 tcttcaacaggtgtagacaataatgaatgatacagaagatgataatacctacaccccttgagac 2220  
QY 2221 aaaaataagcccttaagcaatctcctaattgattccgaagcgaatcggatlaaaatgtagaacctc 2280  
2221 aaaaataagcccttaagcaatctcctaattgattccgaagcgaatcggatlaaaatgtagaacctc 2280  
Db 2221 aaaaataagcccttaagcaatctcctaattgattccgaagcgaatcggatlaaaatgtagaacctc 2280  
QY 2281 gcaacagcagltgtatgacatcgaagaagcgttccaatcttaacgtgaagcggggaagaatg 2340  
2281 gcaacagcagltgtatgacatcgaagaagcgttccaatcttaacgtgaagcggggaagaatg 2340  
Db 2281 gcaacagcagltgtatgacatcgaagaagcgttccaatcttaacgtgaagcggggaagaatg 2340  
QY 2341 ctctcgttaaaaagcggcgaacgaataagttatlaatgacagcgtlccgttcaattgagttt 2400  
2341 ctctcgttaaaaagcggcgaacgaataagttatlaatgacagcgtlccgttcaattgagttt 2400  
Db 2341 ctctcgttaaaaagcggcgaacgaataagttatlaatgacagcgtlccgttcaattgagttt 2400  
QY 2401 gatctgagcaggaaggtcacaataatccagatgttccaactataatagcgggtcaattccac 2460  
2401 gatctgagcaggaaggtcacaataatccagatgttccaactataatagcgggtcaattccac 2460  
Db 2401 gatctgagcaggaaggtcacaataatccagatgttccaactataatagcgggtcaattccac 2460  
QY 2461 ctctcgttctgatacccaagccttccgaatgacagcagctlaatgaaatcaatlaagccactc 2520  
2461 ctctcgttctgatacccaagccttccgaatgacagcagctlaatgaaatcaatlaagccactc 2520  
Db 2461 ctctcgttctgatacccaagccttccgaatgacagcagctlaatgaaatcaatlaagccactc 2520  
QY 2521 caattctctaataatgctccagagcaacagtgccctatgaaatcaatcgttataataa 2580

```

|||||
Db 2521 caatcccaaaatgctccagagaccagctgacctatgaattacccgatacgtatgtaataa 2580
OY 2581 gacaaatggggaataatattctgttataacccaataaactctagcaaccatcaat 2640
    |||||||
Db 2581 gacaaatggggaataatattctgttataacccaataaactctagcaaccatcaat 2640
OY 2641 ttgcccagcgggaataatgggcaatacgtctacagcggtaaggtagaggaatccaccctt 2700
    |||||||
Db 2641 ttgcccagcgggaataatgggcaatacgtctacagcggtaaggtagaggaatccaccctt 2700
OY 2701 ggtcaagcagggggaagggtccaaatgataatcatatcatctcctcacaagg 2760
    |||||||
Db 2701 ggtcaagcagggggaagggtccaaatgataatcatatcatctcctcacaagg 2760
OY 2761 gtaagccagaccagcgtataaagaatagaaaa 2794
    |||||||
Db 2761 gtaagccagaccagcgtataaagaatagaaaa 2794

```

## RESULT 2

US-09-463-238-2

```

: Sequence 2, Application US/09463238
: GENERAL INFORMATION:
: APPLICANT: Edwards, Elizabeth A
: APPLICANT: Smith, Allison M
: APPLICANT: Bustos Guillen, Regla
: APPLICANT: Martin, Catherine R
: APPLICANT: Plant Bioscience Limited
: TITLE OF INVENTION: Starch Debranching Enzymes
: FILE REFERENCE: 97.118
: CURRENT APPLICATION NUMBER: US/09/463,238
: CURRENT FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: PCT/GB98/02280
: PRIOR FILING DATE: 1998-07-30
: PRIOR APPLICATION NUMBER: GB 9716185.5
: PRIOR FILING DATE: 1997-07-31
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 2706
: TYPE: DNA
: ORGANISM: Solanum tuberosum
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2641)
: OTHER INFORMATION: n = a or g or c or t
US-09-463-238-2

```

```

Query Match 1.9%; Score 52; DB 5; Length 2706;
Best Local Similarity 52.9%; Pred. No. 3.6e-05;
Matches 164; Conservative 0; Mismatches 132; Indels 14; Gaps 2;

```

```

OY 1585 ataaaaagtttaaggaatggtcttccatccatcgtgaacacatctggggttaacatg 1644
    |||||
Db 1036 ataaacgaatttaagttatctgttcaaggaacataaagcgtggaatcgtatcatcg 1098
OY 1645 gattgtgtataatcatccttgc-----ccaagcaatctcgtactga 1691
    |||||||
Db 1099 gatgtgttcttaacacacgtctcgaaggaatgaaatggtcccatcattcattaga 1158
OY 1692 taagaattgacagaataattatcaccgtacgtatccaggaatgaattataccaagatca 1751
    |||||
Db 1159 ggcattgacaaacagtggtttatacgtcagctccaaagggtgaatttacaactactca 1218
OY 1752 ggtactggaatgaattgacangaaagccaatggttcaaaaaattatattgactc 1811
    |||||
Db 1219 ggaatggaat-accctcaactgtaataatcccatgtaacgtatcattatagtgatcg 1277
OY 1812 ccttaagttatgggtacatgagatcatatgacggctcgtttgacttaatggcgt 1871
    |||||
Db 1278 ctgagatatgtgggttaaccgaatgacagtagatggcttcgcttgcattcgtctcat 1337

```

```

OY 1872 gcttgaaaa 1881
    |||
Db 1338 ccttaacaga 1347

```

## RESULT 3

US-09-463-238-11

```

: Sequence 11, Application US/09463238
: GENERAL INFORMATION:
: APPLICANT: Edwards, Elizabeth A
: APPLICANT: Smith, Allison M
: APPLICANT: Bustos Guillen, Regla
: APPLICANT: Martin, Catherine R
: APPLICANT: Plant Bioscience Limited
: TITLE OF INVENTION: Starch Debranching Enzymes
: FILE REFERENCE: 97.118
: CURRENT APPLICATION NUMBER: US/09/463,238
: CURRENT FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: PCT/GB98/02280
: PRIOR FILING DATE: 1998-07-30
: PRIOR APPLICATION NUMBER: GB 9716185.5
: PRIOR FILING DATE: 1997-07-31
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 2806
: TYPE: DNA
: ORGANISM: Solanum tuberosum
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (822, 826, 2707 and 2797)
: OTHER INFORMATION: n = a or g or c or t
US-09-463-238-11

```

```

Query Match 1.6%; Score 45.2; DB 5; Length 2806;
Best Local Similarity 54.6%; Pred. No. 0.0034;
Matches 89; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

```

OY 2294 ttatgcccacaaagcgtccatcattcgaagcggggaagaatgctctgtananaag 2353
    |||||
Db 1958 ttatggttcccaagtggtcccaatgatatataatggtcgtatggtacacactaagg 2017
OY 2354 gcgcagacgacaatgatttaaatgctcagtcagtcagtcagtcagtcagtcagtcagtc 2413
    |||||
Db 2018 gaggaaacacaaacagatttgccatgataataataattactcgttgggataaga 2077
OY 2414 aagctcaatataccagatgtttcaactatataagcggtcaat 2456
    |||||
Db 2078 agatgaatctcatctgatttttgagatttgcgacctcat 2120

```

## RESULT 4

US-09-463-238-1

```

: Sequence 1, Application US/09463238
: GENERAL INFORMATION:
: APPLICANT: Edwards, Elizabeth A
: APPLICANT: Smith, Allison M
: APPLICANT: Bustos Guillen, Regla
: APPLICANT: Martin, Catherine R
: APPLICANT: Plant Bioscience Limited
: TITLE OF INVENTION: Starch Debranching Enzymes
: FILE REFERENCE: 97.118
: CURRENT APPLICATION NUMBER: US/09/463,238
: CURRENT FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: PCT/GB98/02280
: PRIOR FILING DATE: 1998-07-30
: PRIOR APPLICATION NUMBER: GB 9716185.5
: PRIOR FILING DATE: 1997-07-31
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1

```

LENGTH: 2843  
TYPE: DNA  
ORGANISM: Solanum tuberosum  
US-09-463-238-1

Query Match 1.4%; Score 39.4; DB 5; Length 2843;  
Best Local Similarity 58.3%; Pred. No. 0.16;  
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 1746 ggtcaggtctcggaaatgaattgcngcgaaggcgaattgtcaaaatttat 1805  
DB 1518 gtatcgaataatcaaaatgaattgcaattaccacatgccaacaaatgattt 1577

OY 1806 tgattcccttaagatattgggtcgaatgatacatattgacggtccggtttgac 1860  
DB 1578 ggaactgtctccgcacacgggttaattgattcatattgattgattttgttc 1632

RESULT 5  
US-09-463-238-10  
Sequence 10, Application US/09463238  
GENERAL INFORMATION:

APPLICANT: Edwards, Elizabeth A  
APPLICANT: Smith, Allison M  
APPLICANT: Bustos Guillen, Regla  
APPLICANT: Martin, Catherine R  
APPLICANT: Plant Bioscience Limited  
TITLE OF INVENTION: Starch Debranching Enzymes  
FILE REFERENCE: 97.118  
CURRENT APPLICATION NUMBER: US/09/463,238  
CURRENT FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: PCT/GB98/02280  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: GB 9716185.5  
PRIOR FILING DATE: 1997-07-31  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 2843  
TYPE: DNA  
ORGANISM: Solanum tuberosum  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1868, 1876)  
OTHER INFORMATION: n = a or g or c or t  
US-09-463-238-10

Query Match 1.4%; Score 39.4; DB 5; Length 2843;  
Best Local Similarity 58.3%; Pred. No. 0.16;  
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 1746 ggtcaggtctcggaaatgaattgcngcgaaggcgaattgtcaaaatttat 1805  
DB 1518 gtatcgaataatcaaaatgaattgcaattaccacatgccaacaaatgattt 1577

OY 1806 tgattcccttaagatattgggtcgaatgatacatattgacggtccggtttgac 1860  
DB 1578 ggaactgtctccgcacacgggttaattgattcatattgattgattttgttc 1632

RESULT 6  
US-09-463-238-3  
Sequence 3, Application US/09463238  
GENERAL INFORMATION:  
APPLICANT: Edwards, Elizabeth A  
APPLICANT: Smith, Allison M  
APPLICANT: Bustos Guillen, Regla  
APPLICANT: Martin, Catherine R  
APPLICANT: Plant Bioscience Limited  
TITLE OF INVENTION: Starch Debranching Enzymes  
FILE REFERENCE: 97.118

CURRENT APPLICATION NUMBER: US/09/463,238  
CURRENT FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: PCT/GB98/02280  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: GB 9716185.5  
PRIOR FILING DATE: 1997-07-31  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2634  
TYPE: DNA  
ORGANISM: Solanum tuberosum  
US-09-463-238-3

Query Match 1.3%; Score 37.4; DB 5; Length 2634;  
Best Local Similarity 67.1%; Pred. No. 0.57;  
Matches 53; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1591 gatttaaggaatggtcttctcactccatcgtgaacacattgggttaacatgatt 1650  
DB 1031 gattcgaagaatggtcgaagccttgcattgctgctgaattgagtcatttagatt 1090  
OY 1651 gttataatcatcattgtg 1669  
DB 1091 gttataatcaccacaatg 1109

RESULT 7  
US-60-248-505-202  
Sequence 202, Application US/60248505  
GENERAL INFORMATION:

APPLICANT: Beasley, Ellen  
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
FILE REFERENCE: c1000918  
CURRENT APPLICATION NUMBER: US/60/248,505  
CURRENT FILING DATE: 2000-11-15  
NUMBER OF SEQ ID NOS: 1998  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 202  
LENGTH: 264546  
TYPE: DNA  
ORGANISM: human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)-(264546)  
OTHER INFORMATION: n = A,T,C or G  
US-60-248-505-202

Query Match 1.3%; Score 35.8; DB 6; Length 264546;  
Best Local Similarity 51.6%; Pred. No. 24;  
Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

OY 1381 aaaggaacaaagggccctgacgaatgaacggtgaatgattccttaaacacttggg 1440  
DB 34030 aaggaagaagaagaataaagaacattgaatggaagaatgaataaacttattg 34089

OY 1441 attactatgttcagcttattgctgttttcgacatcaacagtgctgagtgaaactgaca 1500  
DB 34090 ctggcagatgatattgcttattgcttattcgtaaaatcagagttaccabaaagatagt 34149

OY 1501 accaagaataatgggttatgaccctgcgaacatgat 1539  
DB 34150 aaaaactaataagagttattcagaggttcgacagaagat 34188

RESULT 8  
US-60-248-505-164  
Sequence 164, Application US/60248505  
GENERAL INFORMATION:

```

?
?
? APPLICANT: Beasley, Ellen
? TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
? TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPOR
? TITLE OF INVENTION: PROTEINS, AND USES THEREOF
? FILE REFERENCE: c100918
? CURRENT APPLICATION NUMBER: US/60/248,505
? CURRENT FILING DATE: 2000-11-15
? NUMBER OF SEQ ID NOS: 1998
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 164
? LENGTH: 91051
? TYPE: DNA
? ORGANISM: human
? US-60-248-505-164

```

Query Match	1.3%	Score 35	DB 6	Length 91051
Best Local Similarity	49.7%	Pred. NO. 23		
Matches 89	Conservative 0	Mismatches 90	Indels 0	Gaps 0

Oy 1379 aaaaaggaacaaaaggccctgacaagtaagaagggatgatcctctaanaaacatttg 1438  
| | ||||| ||| | ||| | ||| |||  
Db 89551 agatgggaagaaatttcctccaattcatgtacaatgataatatttcaataactta 89610

Oy 1439 ggattactcagtttcaggcttatgccttglttgcacatacacagtgtcgatgaaatgatc 1498  
| | | | | | | | | | | | | |  
Db 89611 gaatgcctccaagtttaatgcacatcatcttttagatccaaaataaarqgaacyaatgttat 89670

QY 1499 caaccaagaataattggygtatagccctcgcaactatgatcgttcctgaaggcagtac 1557  
|| || | || || || || || || || || || || || || || || || || || || || || ||  
Db 89671 catccatatgtcttttatcttctgtgccttcgggcataaaagattccggaatgttaattac 89729

```

RESULT      9
; Sequence 8, Application US/60248822
; US-60-248-822-8
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINIS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINIS, AND USES THEREOF
; FILE REFERENCE: C1000960
; CURRENT APPLICATION NUMBER: US/60/248,822
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 133937
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(133937)
; OTHER INFORMATION: n = A,T,C or G
; US-60-248-822-8

```

Query Match	1.3%	Score 35	DB 6	Length 133937
Best Local Similarity	55.3%	Pred. No. 28		
Matches 68	Conservative	0	Mismatches	55: Indels 0: Gaps 0

Oy 1542 tccggaaggcaglatcyctacaatgcgaatcgtaatcgataaaaggtttaaga 1601  
| | | | | | | | | | | | | | | |  
Db 120075 ttccggaggttagaagccttgaagcgtaatggttatcaactttacaatagtlttaaca 120134

Oy 1602 aatggttccttcaatccatcgtaaacacattgggtttacaatgatgttgctataatca 1661  
|| ||| | | | ||| || ||| || | || ||  
Db 120135 aaagatttatcttcaaaattacacaccgcgcaaaaatacattgtatttctttaacaca 120194

QY	1662	tac	1664
Db	120195	tac	120197

```

RESULT 10
US-09-262-126C-1/c
: Sequence 1, Application US/09262126C
: GENERAL INFORMATION:
: APPLICANT: Miller, Brian S.
: APPLICANT: Shetty, Jayarama K.
: TITLE OF INVENTION: Modified Forms of Pullulanase
: FILE REFERENCE: GC396-2
: CURRENT APPLICATION NUMBER: US/09/262,126C
: CURRENT FILING DATE: 1999-03-03
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2794
: TYPE: DNA
: ORGANISM: Bacillus deramificans
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(2794)
: OTHER INFORMATION: n = A, T, C, or G
US-09-262-126C-1

```

Query Match	1.2%	Score	34.4	DB	5	Length	2794
Best Local Similarity	45.5%	Pred.	No. 4.3				
Matches 122; Conservative	0	Mismatches	146	Indels	0	Gaps	0

Qy 2399 ttgattgagcagaaagctcaataaccagatgtttccaacattatagcggctaacc 2458  
||||| | |||| | |||| | |||| | ||||  
Db 2666 ttgatgtgccattttcccgctccggcaattgatgtgtctacagttttttttgsgttatna 2607

Oy	2459	acctctgcttgatcaccacgccttcgcgatgagacgtataatgaatacgccacc	2518
Db	2606	ACACAATGATATTTCCCCAATTTGTCTTATTAAACATCATCAGTTAATTCATAGGCCACT	2547

QY 2519 tccattcctaatagtccagagaacacagigtgcctatgaattaactgatacaatgtaaa 2578  
|| || | || || || || || || || || || || || || || || || || || || || || ||  
Db 2546 GTGTTTCCTCGACATTTTAGCAATTGCAGGTGGCTAATGATTTTCAATTAAGCTGTCTCATG 2487

QY 2579 aagcaaatggygaatatcatgttgcattataccctaataacigttagcaaccacca 2638  
|| ||| | | | | | | | | | | | |  
Db 2486 CGGAAGGCTGGGTATCAAGACGAAGGTGATTAGCCCGCTATAACTTGAAAACATCT 2427

QY	2639	atttgcgagcgygaaatgycatcaa	2666
		↑	
Db	2426	GGATATTGAGCTTTCGTCCCAATCAA	2399

RESULT 11  
US-60-248-505-127/c

```

RESULT 11
US-60-248-505-127/C
; Sequence 127, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: G1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 127
; LENGTH: 22788
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22788)
; OTHER INFORMATION: n = A,T,C or G
US-60-248-505-127

```





Query Match	1.28;	Score 34.2;	DB 6;	Length 150471;
Best Local Similarity	53.38;	Prod No 50;		

Best Local Similarity 53.3%; Pred. No. 50;  
Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

1275 aaagaatatagaagatgagtcattctatgaaatgtagtgcgtagctttccattgacc 1334

Db 120735 AAAAAAAAAAGGCTCAATATGCTCTTTAGATATTCAGTATGAATACTGTAATTAAATC 120676

1335 taatcgggtatgaaaataaagggaagtatttgctcttacagaaaaggacaaggg 1394

Db 120675 TGTATTAGAACTGAATTAATTGAATTATTGCTTCATACACATCAGTGAATAATATA 120616

QY 1395 ccctgacaacgtaaa 1409

Db 120615 TCTAGACAACCAAGA 120601

Search completed: April 22, 2001, 09:12:11  
Job time: 8392 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2001, 06:09:20 ; Search time 1905.14 Seconds

(Without alignments)  
214.357 Million cell updates/sec

Title: US-09-262-126c-1

Perfect score:

2794

Sequence: 1 gatgggaacacgacacgacat.....cggtaaaaagtaataagaaaa 2794

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 73081774 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estin5:\*  
68: em\_estom1:\*  
69: em\_estom2:\*  
70: em\_estov1:\*  
71: em\_estov2:\*  
72: em\_estp1:\*  
73: em\_estp12:\*  
74: em\_estp13:\*  
75: em\_estp14:\*  
76: em\_estp15:\*  
77: em\_estp16:\*  
78: em\_estp17:\*  
79: em\_estp18:\*  
80: em\_estp19:\*  
81: em\_estp110:\*  
82: em\_estro1:\*  
83: em\_estro2:\*  
84: em\_estro3:\*  
85: em\_estro4:\*  
86: em\_estro5:\*  
87: em\_estro6:\*  
88: em\_estro7:\*  
89: em\_estro8:\*  
90: em\_estro9:\*  
91: em\_estro10:\*  
92: em\_estro11:\*  
93: em\_estro12:\*  
94: em\_estro13:\*  
95: em\_estro14:\*  
96: em\_estro15:\*  
97: em\_estro16:\*  
98: em\_estro17:\*  
99: em\_estro18:\*  
100: em\_estro19:\*  
101: em\_estro20:\*  
102: gb\_est25:\*  
103: gb\_est26:\*  
104: gb\_est27:\*  
105: gb\_est28:\*  
106: gb\_est29:\*  
107: gb\_est30:\*  
108: gb\_est31:\*  
109: gb\_est32:\*  
110: gb\_est33:\*  
111: gb\_est34:\*  
112: gb\_est35:\*  
113: gb\_est36:\*  
114: gb\_est37:\*  
115: gb\_est38:\*  
116: gb\_est39:\*

117: gb\_est48.\*  
 118: gb\_est49.\*  
 119: gb\_est50.\*  
 120: gb\_est51.\*  
 121: gb\_est52.\*  
 122: gb\_est53.\*  
 123: gb\_est54.\*  
 124: gb\_est55.\*  
 125: gb\_est56.\*  
 126: gb\_est57.\*  
 127: gb\_est58.\*  
 128: gb\_est59.\*  
 129: gb\_est60.\*  
 130: gb\_est61.\*  
 131: gb\_est62.\*  
 132: gb\_est63.\*  
 133: gb\_est64.\*  
 134: gb\_est65.\*  
 135: gb\_est66.\*  
 136: gb\_est67.\*  
 137: gb\_est68.\*  
 138: gb\_est69.\*  
 139: gb\_est70.\*  
 140: gb\_est71.\*  
 141: gb\_est72.\*  
 142: gb\_est73.\*  
 143: gb\_est74.\*  
 144: gb\_est75.\*  
 145: gb\_est76.\*  
 146: gb\_est77.\*  
 147: gb\_est78.\*  
 148: gb\_est79.\*  
 149: gb\_est80.\*  
 150: gb\_est81.\*  
 151: gb\_est82.\*  
 152: gb\_est83.\*  
 153: gb\_est84.\*  
 154: gb\_est85.\*  
 155: gb\_est86.\*  
 156: gb\_est87.\*  
 157: gb\_est88.\*  
 158: gb\_est89.\*  
 159: gb\_est90.\*  
 160: gb\_est91.\*  
 161: gb\_est92.\*  
 162: gb\_est93.\*  
 163: gb\_est94.\*  
 164: gb\_est95.\*  
 165: gb\_est96.\*  
 166: gb\_est97.\*  
 167: gb\_est98.\*  
 168: gb\_est99.\*  
 169: gb\_est100.\*  
 170: gb\_est101.\*  
 171: gb\_est102.\*  
 172: gb\_est103.\*  
 173: gb\_est104.\*  
 174: gb\_est105.\*  
 175: gb\_est106.\*  
 176: gb\_est107.\*  
 177: gb\_est108.\*  
 178: gb\_est109.\*  
 179: gb\_est110.\*  
 180: gb\_est111.\*  
 181: gb\_est112.\*  
 182: gb\_est113.\*  
 183: gb\_est114.\*  
 184: gb\_est115.\*  
 185: gb\_est116.\*  
 186: gb\_est117.\*  
 187: gb\_est118.\*  
 188: gb\_est119.\*  
 189: gb\_est120.\*

190: em\_gss\_pln1.\*  
 191: em\_gss\_pln2.\*  
 192: em\_gss\_pro.\*  
 193: em\_gss\_rod1.\*  
 194: em\_gss\_rod2.\*  
 195: em\_gss\_rod3.\*  
 196: em\_gss\_rod4.\*  
 197: em\_gss\_rod5.\*  
 198: em\_gss\_vrt1.\*  
 199: em\_gss\_vrt2.\*  
 200: em\_gss\_vrt3.\*  
 201: gb\_gss1.\*  
 202: gb\_gss2.\*  
 203: gb\_gss3.\*  
 204: gb\_gss4.\*  
 205: gb\_gss5.\*  
 206: gb\_gss6.\*  
 207: gb\_gss7.\*  
 208: gb\_gss8.\*  
 209: gb\_gss9.\*  
 210: gb\_gss10.\*  
 211: gb\_gss11.\*  
 212: gb\_gss12.\*  
 213: gb\_gss13.\*  
 214: gb\_gss14.\*  
 215: gb\_gss15.\*  
 216: gb\_gss16.\*  
 217: gb\_gss17.\*  
 218: gb\_gss18.\*  
 219: gb\_gss19.\*  
 220: gb\_gss20.\*  
 221: gb\_gss21.\*  
 222: gb\_gss22.\*  
 223: gb\_gss23.\*  
 224: gb\_gss24.\*  
 225: gb\_gss25.\*  
 226: gb\_gss26.\*  
 227: gb\_gss27.\*  
 228: gb\_gss28.\*  
 229: gb\_gss29.\*  
 230: gb\_gss30.\*  
 231: gb\_gss31.\*  
 232: gb\_gss32.\*  
 233: gb\_gss33.\*  
 234: gb\_gss34.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	49.6	1.8	874	119	AW727770 GA_Ea002
2	45.6	1.6	974	125	AZ673433
3	45	1.6	578	167	BE493792
4	44.2	1.6	1201	229	CNS0168L
5	43.4	1.6	1101	229	CNS0039G
6	43	1.5	373	202	AO100007
7	42.8	1.5	558	174	BG227648
8	42	1.5	644	24	A1737641
9	40.8	1.5	612	167	BE444461
10	39.8	1.4	365	118	AW697292
11	39.8	1.4	532	150	BF587295
12	39.8	1.4	921	119	AW730507
13	39.6	1.4	893	223	AZ532406
14	39.6	1.4	920	225	AZ683014
15	39.6	1.4	923	225	AZ668671
16	39.2	1.4	558	120	AW632588
17	39.2	1.4	912	223	AZ549196
18	39	1.4	611	110	AW034518

```
C 19 39 1.4 862 225 AZ685238 ENT1F31TR
C 20 39 1.4 1201 229 CNS0163T
C 21 38.6 1.4 295 147 BF373589 MKO-PT017
C 22 38.6 1.4 304 147 BF373602 MKO-PT017
C 23 38.6 1.4 614 118 AM648801 EST327255
C 24 38.6 1.4 895 223 AZ538451 ENTFR54TR
C 25 38.6 1.4 927 225 AZ677847 ENTFR50TR
C 26 38.6 1.4 1076 142 BE966371 601660419
C 27 38.4 1.4 492 32 AM010860
C 28 38.4 1.4 578 204 AO280563 C1TBT-E1-
C 29 38.4 1.4 612 204 AO280573 C1TBT-E1-
C 30 38.4 1.4 736 213 AO280573 C1TBT-E1-
C 31 38.4 1.4 864 225 AZ676090 ENT1Q90TR
C 32 38.4 1.4 884 225 AZ680209 ENTMB60TR
C 33 38.4 1.4 897 223 AZ528531 ENTBM01TR
C 34 38.4 1.4 934 225 AZ673277 ENTBA10TR
C 35 38.2 1.4 676 116 AM506191 GP1528 G1
C 36 38.2 1.4 900 140 BE870915 601448867
C 37 37.6 1.3 350 138 BE681758 179528 MA
C 38 37.6 1.3 569 231 CNS03FGA AL241651 Tetradon
C 39 37.6 1.3 602 205 AO380419 RPI11-15
C 40 37.6 1.3 1123 158 W29506 mc03f04.r1
C 41 37.4 1.3 677 23 A1653243
C 42 37.4 1.3 749 210 AO740635 HS-5508-A
C 43 37.4 1.3 914 173 BG105065 602312349
C 44 37.4 1.3 946 232 CNS056JM Tetradon
C 45 37.2 1.3 1101 229 CNS0060A AL065616 Drosophila
```

## ALIGNMENTS

```
RESULT 1
LOCUS AM727770 874 bp mRNA EST 16-NOV-2000
DEFINITION GA_Ea0028A17 Gossypium arboreum 7-10 dpa fiber library Gossypium
VERSION AM727770
KEYWORDS
SOURCE
ORGANISM Gossypium arboreum.
REFERENCE Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Maltaves; Malvaceae; Gossypium.
1 (bases 1 to 874)
Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
High quality sequence stop: 874.
FEATURES
Source
Location/Qualifiers
1..874
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0028A17"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pAK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 238 a 148 c 225 g 261 t 2 others
ORIGIN
```

```
Query Match 1.8%; Score 49.6; DB 119; Length 874;
Best Local Similarity 50.9%; Pred. No. 0.0052;
Matches 177; Conservative 0; Mismatches 156; Indels 15; Gaps 2;
OY 1555 tatgtacaatgacgaatgtaatgctgataaagaagtttaagaaatgcttca 1614
DB 25 TATTCAGAGATCAAAATGTCCTGCGCCATATGATTTAGAAATGATTCAGGCA 84
OY 1615 ctccatcgtaacacatggtggttaacatgtagtctgctataacatcccttcacg 1674
DB 85 CTCACCCATATGTCCTGCTGTTGTGTGATGTCGTATTAATCAATCTGCATGCMAAT 144
OY 1675 c-----aaatcctgacttcgataaattgtaacgaatattatccgtacg 1722
DB 145 GGGCCTTTGATTAAGATTTCTCTCTGATTAAGATTTGTCAGGTTACTATCTGAGAAC 204
OY 1723 atgacccagtaattataccaaagatcagtagctggaatgaaatlgcangcngaaag 1782
DB 205 AACACTGATGCTTTATTCAGAACAGT---ACATGCTTAACAAATACGTCTAGTGACAT 261
OY 1783 ccaatggtcaaaaattatattatgattcccttaagtattggtcaatgataatatt 1842
DB 262 TATATGTCGACCGCTTATTTGATGATCTTTGAGCTGGCAATAATATTAAGTT 321
OY 1843 gacgcttcggttctgacttaagtcgctgctgtgtaaaagacagatg 1890
DB 322 GATGGCTTTCGTTTGTGATCTCATGGCTCACATCATCAACGACGATG 369
```

## RESULT 2

```
AZ673433 974 bp DNA GSS 14-DEC-2000
LOCUS ENT1K90TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION ENT1K90TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
VERSION AZ673433
KEYWORDS genomic, DNA sequence.
SOURCE
ORGANISM Entamoeba histolytica.
REFERENCE Eukaryota; Entamoebidae; Entamoeba.
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
1 (bases 1 to 974)
Determination of clone end sequences from Entamoeba histolytica
TITLE HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seg primer: M13-Reverse
Class: shotgun
High quality sequence start: 32
High quality sequence stop: 182.
FEATURES
Source
Location/Qualifiers
1..974
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
```





High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu

Sequence Tagged Connector  
Plate: 3054 row: L column: 13  
Class: BAC ends  
High quality sequence stop: 373.  
Location/Qualifiers

# FEATURES

## SOURCE

1. 373  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-3054 Col-13 Row=L"  
/clone\_1lb="C1T Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBeloBAC11, BAC Clones in E-coli DH10B"  
BASE COUNT 150 a 36 c 67 g 120 t  
ORIGIN

Query Match 1.58; Score 43; DB 202; Length 373;

Best Local Similarity 50.28; Pred. No. 0.31; Mismatches 105; Indels 0; Gaps 0;

Matches 106; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1260 taacacataatcgcacaaagataagagatgcatcctgaatgagtcgtcgtga 1319

DB 144 TTAATATTACCTCCAAAGTACTATTAAGCTTGATTAATTAATTAATTAATTAAT 203

QY 1320 ctttcacatgacccatcctcgtggtatgagaataaataagagagatctgctctacaga 1379

DB 204 AATAGATATTAATCACTTACTTACCTATCAGACACAGAGTTGAAGAGAGGAGAAATA 263

QY 1380 aaagagaaacaaagccctgcacacgtlaaagagcgagatagatccttaaacacacttg 1439

DB 264 AGAAGGAAGAAAGAGAGATATTAAGAAAGAAAGAAAGAAAGAAAGAGATGT 323

QY 1440 gattactcatgtctcagctcgtcttcttc 1470

DB 324 TATGCTAGCATGCTGTTTTTTTATTTTTC 354

## RESULT 7

BG227648 558 bp mRNA EST 08-FEB-2001

LOCUS kq13e10.y1 TBN95TM-SSR Strongyloides stercoralis cDNA 5' similar to

ACCESSION WP:110F2.1 CE02040 GLCYL-TRNA SYNTHETASE ;, mRNA sequence.

VERSION BG227648.1 GI:12715282

KEYWORDS EST.

SOURCE Strongyloides stercoralis.

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Strongyloides stercoralis; Strongyloides; Strongyloides.

REFERENCE 1 (bases 1 to 558)

AUTHORS McCarter,J., Clifton,S., Marra,M., Hillier,L., Kucaba,T., Martin,J.

Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen

M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey

N., Schurk,R., Ralder,E., Kohn,S., Shin,T., Jackson,Y., Cardenas

M., McCann,R., Waterston,R. and Wilson,R.

THE WASHINGTON UNIV. Nematode EST Project, 1999

Unpublished (1999)

CONTACT: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

The library was constructed by Dr. Thomas Nutman and colleagues of

NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington

University Genome Sequencing Center St. Louis.

High quality sequence stop: 395.

Location/Qualifiers

1. 558

/organism="Strongyloides stercoralis"

/strain="Rhabditiform larvae obtained from gerbils"

/db\_xref="taxon:6248"

/clone\_1lb="TBN95TM-SSR"

/lab\_host="XL-1 Blue MRF" (Stratagene); Site:1:

/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site:1:

EcORI; Site:2: XhoI; mRNA was purified from 2 x 10E3

Rhabditiform larvae which had been isolated from gerbils

experimentally infected with larvae originally isolated

from experimentally infected dogs. cDNA was constructed

and, using adaptors, was cloned unidirectionally into the

vector from the EcoRI site to the XhoI site. The library

has an unamplified titer of 1 x 10E5 pfu/ml and an

amplified, undiluted titer of 9 x 10E11 pfu/ml. The

average insert size of the unamplified library is 675 bp

(range, 100-1700)."

BASE COUNT 205 a 71 c 100 g 182 t

ORIGIN

Query Match 1.58; Score 42.8; DB 174; Length 558;

Best Local Similarity 47.78; Pred. No. 0.41;

Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 1490 aaactgattccacccacagataatggtgtatgacccctgcgaactatgagtcctgaag 1549

DB 17 AAGATGTCAAAACAGGTGAATGTTTGTGCGCCATTTAATCAAAATATCTGTA 76

QY 1550 ggcagatgctcacaatgcaatgcaatgctcgtatataaagagcttaagaaatgctc 1609

DB 77 AACTTATATCTGTAATTAATAAACTAGTACTGTAATTAATAAGACGACTAATGACGCTTTAA 136

QY 1610 ttctaccatcgcgtgaacacactgggttaacatgagatgctgatacatcatccttg 1669

DB 137 ATAGACTTGATGCTTTGATATATATGATATTTCAAAATATTTATTAATAAGTACAACTTTA 196

QY 1670 ccaagcaaatcctcgtactcgtacataaattgacagaaatattacacgtacgtagtc 1729

DB 197 AGTACCACTTACGCGCATATTTAATCTGACCAACAAAGCTTTAATCTCATGTTTCCAA 256

QY 1730 aggtcaattatacacaagatca 1751

DB 257 CTGAATTTGACACCAACCGCTGA 278

RESULT 8

A1737641/c 644 bp mRNA EST 02-FEB-2000

LOCUS 605036B10.x2 605 - Endospore cDNA library from Schmidl lab Zea mays

ACCESSION A1737641

VERSION A1737641.1 GI:5069676

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae

; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 644)

AUTHORS Walbot,V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

CONTACT: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221



Email: walbot@stanford.edu  
 Plate: 605036 Row: B Column: 10.  
 Location/Qualifiers  
 source 1. .644

/organism="Zea mays"  
 /cultivar="Ohio43"  
 /db.xref="taxon:4577"  
 /clone\_lib="605 - Endosperm cDNA library from Schmidt lab"  
 /tissue\_type="nucellar, embryo, and endosperm"  
 /dev\_stage="10-14 days post-pollination"  
 /lab\_host="DH5(alpha)"  
 /note="Organ: kernel; Vector: PAD-GAL4-2; Site:1: EcORI;  
 Site\_2: XhoI; kernel endosperm cDNA library from Schmidt  
 lab"

BASE COUNT 165 a 162 c 161 g 156 t  
 ORIGIN

Query Match 1.58; Score 42; DB 24; Length 644;  
 Best Local Similarity 54.2%; Pred. No. 0.75;  
 Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Oy 2302 tcacagagcgttcattcgcacgaagcgaggaatgcttcgtanaaagcgagcaac 2361  
 |||||  
 Db 640 TCTCAGGAGTTCATGTTCTCATGCGCATGATATGTCACACAAAGGAGGCAAC 581  
 Oy 2362 gacatagttataatgcagcgatgcgtcgaatgagttgattgagcaggaagctcaa 2421  
 |||||  
 Db 580 AACATAGTACTGCGCATGACCATATATGTCATTTTCCTGGATTAAGAGAGAGA 521  
 Oy 2422 tatccagatgtttcaactatattatagcgggctaat 2456  
 |||||  
 Db 520 CAATCCTCTGATTGTACAGATTCTGCGGTCTCAT 486

RESULT 9  
 BE444461 612 bp mRNA EST 25-JUL-2000  
 LOCUS WHE1111\_H10\_P1925 Wheat etiolated seedling root normalized cDNA  
 DEFINITION library Triticum aestivum cDNA clone WHE1111\_H10\_P19, mRNA  
 sequence.

ACCESSION BE444461.1 GI:9444009  
 VERSION BE444461  
 KEYWORDS EST.  
 SOURCE bread wheat.  
 ORGANISM Triticum aestivum  
 Magnoliophyta: Liliopsida; Poales; Poaceae; Pooidae; Triticeae;  
 Triticum.  
 1 (bases 1 to 612)

REFERENCE  
 AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han  
 P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,  
 Rausch,C.J., Seaton,C.L., Tong,J.C. and Zhang,D.  
 The structure and function of the expressed portion of the wheat  
 genomes - Normalized root cDNA library  
 unpublished (2000)

JOURNAL  
 COMMENT US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105959773  
 Fax: 5105959818  
 Email: canderan@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Stragene SK primer.  
 Location/Qualifiers

FEATURES  
 Source

1. .612  
 /organism="Triticum aestivum"  
 /cultivar="Chinese Spring"  
 /db.xref="taxon:4565"  
 /clone="WHE1111\_H10\_P19"  
 /clone\_lib="Wheat etiolated seedling root normalized cDNA"

library"  
 /tissue\_type="Root"  
 /dev\_stage="Five day old etiolated seedling"  
 /lab\_host="E. coli DH10s"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid  
 phagescript SK; Site\_1: EcORI; Site\_2: XhoI; Seeds were  
 surface-sterilized, germinated and grown aseptically in  
 the dark at room temperature on filter paper with water,  
 nystatin and cefotaxime in covered crystallization  
 dishes. Roots were harvested. The tissue, total RNA, and  
 poly(A) RNA were prepared, a cDNA library was made in the  
 T3 close lab (Choi, Close, Fenton) at the university of  
 California, Riverside. The cDNA clones were in vivo  
 normalized to give phagescript phagemids before  
 normalization was carried out. The mass excision of  
 phagemid library and normalization were done in HT Nguyen  
 lab by D. Zhang at Texas Tech University. Normalization  
 protocol used was that of Soares. Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

BASE COUNT 175 a 130 c 128 g 179 t  
 ORIGIN

Query Match 1.5%; Score 40.8; DB 167; Length 612;  
 Best Local Similarity 48.8%; Pred. No. 1.6;  
 Matches 139; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

Oy 2152 gacttacctcttcaccaggtgagacaataactatgtacaaagtcataactaac 2211  
 |||||  
 Db 8 GCGTATACCTTCCTCCCGATGAGAAATTAATTAATGTTCTGCTCATGACATGAGCT 67  
 Oy 2212 ctttggacacaatagccctaaagcaatc---taatgattccgaagcgagtgattaa 2268  
 |||||  
 Db 68 CTGTTGATGATTATCAGTTAAGACCCCAATGMACTTTCAGTTGATAGAGATCAG 127  
 Oy 2269 atgagatgaactgcacaaagcagttgattagctcaaaagcgcttcattcagcag 2328  
 |||||  
 Db 128 ATAAATCATTTGGCCCTCCAGCATGATGCGATTTCGCCAGGAATACCCCTTTCATGCT 187  
 Oy 2329 ggggaagaaatgcttcgtanaaagcgagcaagacaaatagttataatgcagcgatg 2388  
 |||||  
 Db 188 GGTGACGAGACTAATGAATCTAATGTCATTTGATGATTCATACAACTCTGCTGATGG 247  
 Oy 2389 gtcaatgaattgattgagcaggaagctcaatcagatggt 2433  
 |||||  
 Db 248 TTTAACAACTGATTGTTTACCTATGAAACAAACAAATGGGCTT 292

RESULT 10  
 AM697292 365 bp mRNA EST 21-DEC-2000  
 LOCUS NF117C04ST1F1033 Developing stem Medicago truncatula cDNA clone  
 DEFINITION NF117C04ST 5', mRNA sequence.  
 ACCESSION AM697292  
 VERSION AM697292  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 Fabales; Fabaceae; Papilionoideae; Medicago.  
 1 (bases 1 to 365)

REFERENCE  
 AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
 C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon  
 R.A.

JOURNAL  
 COMMENT Medicago truncatula stem library  
 Unpublished (2000)  
 On Apr 14, 2000 this sequence version replaced gi:7572054.  
 Contact: Dixon RA  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation





Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

BASE COUNT  
ORIGIN

238 a 185 c 81 g 416 t

# Query Match

1.4% Score 39.6; DB 225; Length 920;  
Best Local Similarity 52.4%; Pred. No. 4.3;

Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1222 gacctggctaaacagatcctgcctgctggaacagatgataacatatcagccaaagaat 1281

DB 291 GAAATGCTTCAAAAAGCAGCTTCAATGATGACGTATTATTAAGAATGACAAAATT 232

QY 1282 atagaagatgaggtcactctatgaatgagtgccgtgactttccatcgacctaatcg 1341

DB 231 ACAATATATGGCGTTCAATGTAAATTTGAATTAATGCTTTCTTCAATCAAGCAAGCAA 172

QY 1342 ggtatgaaaataaagggaagatcttgctctctacagaaaaaggaa 1387

DB 171 GTATTACAACTAAAGAAAATTAATTAAGTCTTATGCAAAATTCTA 126

## RESULT 15

AZ668671

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ668671 923 bp DNA GSS 14-DEC-2000  
ENTJF177P Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

AZ668671  
GSS.  
GI:11805817

Entamoeba histolytica.

Entamoeba histolytica.

Eukaryota; Entamoebidae; Entamoeba.

1 (bases 1 to 923)  
Loftus B., Van Aken S. and Fraser C.  
Determination of clone end sequences from Entamoeba histolytica

HM1:IMSS sheared DNA library  
Unpublished (2000)

Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Forward

Class: shotgun  
High quality sequence start: 24  
High quality sequence stop: 765.

Location/Qualifiers

1..923  
/organism="Entamoeba histolytica"

/strain="HM1:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broch cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a

light size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

BASE COUNT 384 a 113 c 104 g 322 t

ORIGIN

113 c 104 g 322 t

Query Match 1.4% Score 39.6; DB 225; Length 923;  
Best Local Similarity 52.4%; Pred. No. 4.3;  
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1222 gacctggctaaacagatcctgcctgctggaacagatgataacatatcagccaaagaat 1281

DB 56 GAAATGCTTCAAAAAGCAGCTTCAATGATGACGTATTATTAAGAATGACAAAATT 115

QY 1282 atagaagatgaggtcactctatgaatgagtgccgtgactttccatcgacctaatcg 1341

DB 116 ACAATATATGGCGTTCAATGTAAATTTGAATTAATGCTTTCTTCAATCAAGCAAGCAA 175

QY 1342 ggtatgaaaataaagggaagatcttgctctctacagaaaaaggaa 1387

DB 176 GTATTACAACTAAAGAAAATTAATTAAGTCTTATGCAAAATTCTA 221

Search completed: April 22, 2001, 06:52:16  
Job time: 2576 sec

---





PR 15-JUL-1993: 93BE-0000744.  
 PR 19-NOV-1993: 93BE-0001278.  
 XX (SOLV ) SOLVAY SA.  
 PA Amory A, DeMeer P;  
 PI MPI: 1994-210291/26.  
 DR N-PSDB; Q68699.  
 XX  
 PT New pullulanase from *Bacillus deramificans* - for starch  
 PT saccharification etc., with good stability over wide temp. and pH  
 PT ranges, also related DNA vectors, transformed cells etc.  
 XX  
 PS Claim 4; Fig 5; 61pp; French.  
 XX  
 CC This sequence represents the pullulanase enzyme precursor isolated  
 CC from *Bacillus deramificans* T 89.117D. The pullulanase hydrolases  
 CC alpha-1,6-glucosidic bonds of amylopectin and of pullulan. The  
 CC enzyme can be used (with glucoamylase) for saccharification of  
 CC starch. Typical applications include use as anti-staling additives  
 CC in bread making and brewing, in preparation of low calorie foods  
 CC and to clarify fruit juices.  
 CC  
 XX  
 SQ Sequence 957 AA;

Query Match 99.68; Score 5040.5; DB 15; Length 957;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 955; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAKRLIYVCLSVCLVLMFAFNKVGOSAHADGNTTIVHYFCFPAQGVQPMKMMRPDGG 60  
 DB 1 MAKRLIYVCLSVCLVLMFAFNKVGOSAHADGNTTIVHYFCFPAQGVQPMKMMRPDGG 60  
 QY 61 GAERYENQPADSTGAVASADIPGNPQVGIIVTODTKDVSADRYTDLKSGNEVWLVG 120  
 DB 61 GAERYENQPADSTGAVASADIPGNPQVGIIVTODTKDVSADRYTDLKSGNEVWLVG 120  
 QY 121 NSQIFNEKAEADAKRAVSNATLIDASNOLVLRISQPLTIGEGXSGFTVHDTPANKDIPV 180  
 DB 121 NSQIFNEKAEADAKRAVSNATLIDASNOLVLRISQPLTIGEGXSGFTVHDTPANKDIPV 180  
 QY 181 TSVKDAISLGGDVTAVLAGTQFIHIFGSGDMAPDNHSTLTKVTNNLVQESDLEPGNVOYK 240  
 DB 181 TSVKDAISLGGDVTAVLAGTQFIHIFGSGDMAPDNHSTLTKVTNNLVQESDLEPGNVOYK 240  
 QY 241 VALNDSKNN-SYPSDNINLTVPAGAHVTFSTYIPSTHAYVDTINPNADQVSEGVKTL 299  
 DB 241 VALNDSKNN-SYPSDNINLTVPAGAHVTFSTYIPSTHAYVDTINPNADQVSEGVKTL 299  
 QY 300 VVTVLCEDDPVSHLTSTQTDGYAKOVIIPRNVLNNSQYVYSGCDLGNTRYOKATTKVMA 359  
 DB 300 VVTVLCEDDPVSHLTSTQTDGYAKOVIIPRNVLNNSQYVYSGCDLGNTRYOKATTKVMA 359  
 QY 360 PTSTQVNVLLYDSATGSVTKIVPMTASGHWMEATVQNNENNYMYEYVGGOSTRTAVD 419  
 DB 360 PTSTQVNVLLYDSATGSVTKIVPMTASGHWMEATVQNNENNYMYEYVGGOSTRTAVD 419  
 QY 420 PYATATAPNGSTRGMIVDLAKTDPAGMNSDKHITPKNIEDEVIYEMDVRFESIDPNSGMKN 479  
 DB 420 PYATATAPNGSTRGMIVDLAKTDPAGMNSDKHITPKNIEDEVIYEMDVRFESIDPNSGMKN 479  
 QY 480 KCKYLLATEKCTGCPDNVKTGIDSLKQICHTHQMLPVPFASNSVDEIDPQDMMWCTDPKN 539  
 DB 480 KCKYLLATEKCTGCPDNVKTGIDSLKQICHTHQMLPVPFASNSVDEIDPQDMMWCTDPKN 539  
 QY 540 YDVPESQYATNANGNARIKEFEKEMVLSLREHIGVNMDDVYNTTFAQTSDPKIYPEY 599  
 DB 540 YDVPESQYATNANGNARIKEFEKEMVLSLREHIGVNMDDVYNTTFAQTSDPKIYPEY 599  
 QY 600 YRTMIOVITPTDQVLEMLKLAERPNOVKFTISLKYVNEHYHIDGRFDMALLGKDTMS 659  
 DB 600 YRTMIOVITPTDQVLEMLKLAERPNOVKFTISLKYVNEHYHIDGRFDMALLGKDTMS 659

DB 601 YRTMIOVITPTDQVLEMLKLAERPNOVKFTISLKYVNEHYHIDGRFDMALLGKDTMS 660  
 QY 660 KAASELHA1NPGIALYGEPMWGTSALEPDDQLTKGAOKGCMGVAFVNDNLNADGNVFD 719  
 DB 661 KAASELHA1NPGIALYGEPMWGTSALEPDDQLTKGAOKGCMGVAFVNDNLNADGNVFD 720  
 QY 720 SSAQGFATGATGLTDAIKNGVEGSINDFTSSPGETINVTSHDNTTLMDKIALSNPDSE 779  
 DB 721 SSAQGFATGATGLTDAIKNGVEGSINDFTSSPGETINVTSHDNTTLMDKIALSNPDSE 780  
 QY 780 ADRIKMELAOAVVMTSQGVFPNMGGEMLRXKCGNNSYNAGAVNEFPMRSRAOVPDV 839  
 DB 781 ADRIKMELAOAVVMTSQGVFPNMGGEMLRXKCGNNSYNAGAVNEFPMRSRAOVPDV 840  
 QY 840 FNYVSGLIHLRLDHPAPRMTTANESHLQPLNSPENTVAEYLDHVNKDKMGNITVYN 899  
 DB 841 FNYVSGLIHLRLDHPAPRMTTANESHLQPLNSPENTVAEYLDHVNKDKMGNITVYN 900  
 QY 900 PNKTVAITNLPSCGMALNATNSGKVGESTLGOAESVOPGISMILHQEVSPDHGKK 956  
 DB 901 PNKTVAITNLPSCGMALNATNSGKVGESTLGOAESVOPGISMILHQEVSPDHGKK 957

## RESULT 2

Y31746  
 ID Y31746 standard; Protein; 928 AA.

XX Y31746;  
 AC Y31746;  
 DT 22-NOV-1999 (first entry)

DE *Bacillus deramificans* mature pullulanase.  
 XX  
 KW Pullulanase; enzyme engineering; starch; saccharification;  
 KW straight chain amylase; alpha-dextrin 6-glucano-hydrolase;  
 KW alpha-1,6-glucosidase.  
 OS *Bacillus deramificans*.

Key Location/Qualifiers  
 FT MISC-difference 134  
 FT MISC-difference 135  
 FT MISC-difference 135  
 FT MISC-difference 591  
 FT MISC-difference 592  
 FT MISC-difference 783  
 FT MISC-difference 783  
 FT Protein  
 FT 99..928  
 FT /note- "N-truncated pullulanase used in claimed  
 FT composition of Claim 31 (see Y31747)."  
 FT Protein  
 FT 103..928  
 FT /note- "N-truncated pullulanase used in claimed  
 FT composition of Claim 32 (see Y31748)."

MO945124-A2.  
 PD 10-SEP-1999.  
 PF 03-MAR-1999: 99MO-US04627.  
 PR 04-MAR-1998: 98US-0034630.  
 PA (GENV ) GENENCOR INT INC.

PI Miller BS, Shetty JK;  
 XX  
 DR MPI: 1999-540851/45.  
 DR N-PSDB; X87941.

PT New modified pullulanase for saccharification of aqueous liquefied



PT starch  
 XX  
 PS Claim 31; Fig 1A-E; 49pp; English.  
 XX  
 CC The present sequence represents *Bacillus deramificans* strain  
 CC T89.117D mature pullulanase (EC-3.2.1.41). This enzyme has a pI of  
 CC 4.1-4.5, has a half-life of about 55 hr at 60 deg C and is capable  
 CC of catalysing the hydrolysis of alpha-1,6-glucosidic bonds in  
 CC amylopectin and pullulan. The invention relates to the discovery  
 CC that *B. deramificans* pullulanase recombinantly produced in a  
 CC *Bacillus* host is modified yet retains the ability to catalyse the  
 CC hydrolysis of an alpha-1,6-glucosidic bond. The modification of  
 CC the recombinant pullulanase appears to be a result of misprocessing  
 CC of the signal sequence by a signal peptidase as well as  
 CC susceptibility to extracellular proteolytic processing. The  
 CC invention provides modified pullulanases, especially N-terminally  
 CC truncated *B. deramificans* modified pullulanase (see Y31747-48),  
 CC nucleic acids encoding them, vectors, host cells and methods for  
 CC recombinant production. The modified pullulanases are useful in  
 CC the starch industry for the saccharification of aqueous liquified  
 CC starch, forming syrups of high glucose and maltose content. Other  
 CC uses (not claimed) are: as anti-staling additives in bread; in  
 CC brewing low calorie beers; production of low calorie foods (where  
 CC amylose is used as substitute for fat); to clarify fruit juices;  
 CC in the preparation of oligosaccharides from amylopectin or of  
 CC tetraoligosides from maltose; to condense mono- or oligosaccharides,  
 CC or in the pharmaceutical industry (no details). When used for  
 CC starch saccharification, the modified pullulanases reduce the  
 CC amounts of by-products, especially they provide higher glucose  
 CC yields in short reaction times, without losses associated with  
 CC reversion reaction products. When used with glucamylase, they  
 CC make possible a higher level of dissolved solids in the material  
 CC being treated, resulting in increased productivity and a more  
 CC energy-efficient process.  
 CC  
 XX  
 XX Sequence 928 AA;  
 S0  
 Query Match 96.7%; Score 4891.5; DB 20; Length 928;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 927; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 509 ITHVOLMPVFNASVDETDPTQDNMGYDPRNYDVEGQYATNANGNARIKEFEKEMVLSH 568  
 DB 481 lthvqlmpvfaensvdeidcpqdwgydprnydvegqyatanangnarikfekeemvlsjh 540  
 QY 569 RHHIGVNDVYNNHFFATQISDFDKIVPEYRTYRTMIOVLIPIPDQVLEMKLAERPVMQKF 628  
 DB 541 rehigvnmvnyhftatqisdfdkivpeyyrtmlylripdqylemklaerpmvqkf 600  
 QY 629 IIDSILKYWNEVHIDGFRDRLMALLGKDTMSKAASELHAINFGIALYGPMPGTSALPD 688  
 DB 601 iidsilkywnevhiidgfrdrlmalilgkdtmskaaselhainpfdialygpmpgtsalpd 660  
 QY 688 DDLTRTGAOKGKGVAVFNDNLNRLDGNVFDSSAOGFATGATGLDAIKNGEGSINDPT 748  
 DB 661 ddltrtgaokgkvavfndnlrnldgnvfdssagfatgatglcdaikngegsindt 720  
 QY 749 SSPGETINTVISHDNTTMDKTALSNPNDESDRIKMDLAQAVVMTSGVPPMOGGEEM 808  
 DB 721 sspgetintvishdnttmdktalsnpsndeadrlkmdelaqavvmtsgvpfmgggeom 780  
 QY 809 LXXKGGNDVSNAGDAVNEFDMRSKAQVDPVFNYSGLHLRLDHPAPMTTANETNSHL 868  
 DB 781 lxxkgsndvsnagdavnefdmrskaqydpvfnysglhlrlidpafmtttanetnshl 840  
 QY 869 QFLNSPENTVAYELTDHVNKDKMGNIIVYNNPKTVATINLPSGKMAINATSGKVGESTL 928  
 DB 841 qflnspentvayeltdhvnkdkmgniivynnpkvtacatinlpsgkwaيناتsgkvgestl 900  
 QY 928 GQAGSVQVPGISMMILHOEVSPDHGKK 956  
 DB 901 gqaegsvqvgismmilhoevspdhgkx 928

RESULT 3  
 R56989  
 ID R56989 standard; Protein: 928 AA.  
 XX  
 AC R56989;  
 XX  
 DT 20-FEB-1995 (first entry)  
 XX  
 DE *Bacillus deramificans* mature pullulanase.  
 XX  
 KW Pullulanase; recombinant production; pullulan; amylopectin;  
 KM hydrolysis; alpha-1,6-glucosidic bond; starch saccharification.  
 XX  
 OS *Bacillus deramificans* T 89.117D (LMG P-13056).  
 XX  
 FH Key  
 FT Misc-difference 134 Location/Qualifiers  
 FT Misc-difference 135 /note= "corresponds to CGN codon"  
 FT Misc-difference 135 /note= "corresponds to NNA codon"  
 FT Misc-difference 591 /note= "corresponds to CAN codon"  
 FT Misc-difference 592 /note= "corresponds to CGN codon"  
 FT Misc-difference 783 /note= "corresponds to ANA codon"  
 FT  
 XX  
 PN EP605040-A.  
 XX  
 PD 06-JUL-1994.  
 XX  
 XX 20-DEC-1993; 93BP-0203593.  
 PF  
 XX 28-DEC-1992; 92BE-0001156.  
 PR 15-JUL-1993; 93BE-0000744.  
 PR 19-NOV-1993; 93BE-0001276.  
 XX  
 PA (SOLV ) SOLVAT SA.  
 XX

PI Amory A, DeWeer P;  
 XX WPI; 1994-210291/26.  
 DR N-PSDB; Q68698.  
 XX  
 PT New pullulanase from *Bacillus deramificans* - for starch  
 PT saccharification etc., with good stability over wide temp. and pH  
 PT ranges, also related DNA vectors, transformed cells etc.  
 XX  
 PS Claim 3; Fig 4; 61pp; French.  
 XX  
 CC This sequence represents the mature pullulanase enzyme isolated  
 CC from *Bacillus deramificans* T 89.117D. The pullulanase hydrolyses  
 CC alpha-1,6-glycosidic bonds of amylopectin and of pullulan. The  
 CC enzyme can be used (with glucoamylase) for saccharification of  
 CC starch. Typical applications include use as anti-staling additives  
 CC in bread making and brewing, in preparation of low calorie foods  
 CC and to clarify fruit juices.  
 CC  
 XX  
 XX Sequence 928 AA;

Query Match 96.6%; Score 4888.5; DB 15; Length 928;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 926; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 30 DCAHTTIVHFCPCADYQWMSLMMMPKDCGAEYDFNQPADSGAASADIPGNPSQVC 89  
 1 dnttttltivhfcpcadgqpwslmmwpkdgsgydfnqpadesfgavasadiipnpsqvg 60  
 90 IIVRTQWMTQVSDRYTIDSKGNEVWLVEGNSQIFENKDAEPAAPVSNATLDAASNO 149  
 61 lrvrtqwtkvdsadryidtskgnvwlvegnsqiftnkdaedaapvsnaylidsnq 120  
 150 VLVKLSQPLTLEGSCSCTVTHDDTANKDIPVTSYKDAISLGDVYAVIAGTFOHIFGCSDW 209  
 121 vlklsqpltleqsgxgftvhdtdankdipvtavkdaslgdvavlaqlfqbhifgsgdw 180  
 210 APDNHSLTLKKVNNLKYQSGDLPEGNYQYKVALINDSMNN-STPSDINILTVPAAGAHVT 268  
 181 apdnhsellkkvnnlyqsgdlpegnqykvalindsmnpspsdlnltvpagahvt 240  
 269 FSVYPTSHAYDTINNNAALQVESGVKTDLYMTLGEDDVSHTLSIQDGYAKQVIR 328  
 241 fsvyptshaydttinnnadlqvsgvktolvtlvgedpvshtlsiqdgyakqvir 300  
 329 RNVLNSQYVYSGDDLGNVTYTKATPFKVAPESTOVNVLVYSATGSVYKIVPMTASGH 388  
 301 rnvlnsqyvygddlgnvtlqkeltkvwapsltqynvllysatgsvklivpmtasgh 360  
 389 GWEAATYNOULENKYTYEVTGCGSTRTAVDPYATAIAPNGTROMIVDLAKTPDAGNSD 448  
 361 gweaatynqlenwytyevtgsgstrtavyataiapngtrgmivdlaktpdagsnd 420  
 449 KHTTPKIEDEVITKEMVRFSTIDPNSGMKNKGYLALTEKGTGSPNVNTGIDSLKOLG 508  
 421 khttpkiedevitkempvrfstidpnsgmknkgyalalteqgtgspnvntgidslkolg 480  
 509 ITHVOLMPFASNSVDETDPTODNMGYDPFRNYDPECOVATNANGNARIIEFKEMVLSLH 568  
 481 lthvolmpvfasnsvdetdptodnmgydprnydpeqyvalnangnarlefkemvlslh 540  
 569 REHICVMMDVYVNHFTATQISDPDKLVPEYRYKTMIOVITPQVLEHMLKAEHPWQKF 628  
 541 rehivmmdvvyvnhftatqisdpdfklvpeyrytmmqvllptdqvlomklxaeprwqkf 600  
 629 IIDSLEKVMVEYHIDGRFLMLLKGDTMSKASELHAIINPGIALYGEPTGTSALPD 688  
 601 iidslkyvmveyhldgrflmlkgdtkmskaseihainpgialygeptgtsalspd 660  
 689 DQLLTKGACGKGVAVFNDMLRNALDGNVFDSSAOCFATGATGTTDAIKKGVESINDFT 748  
 661 dqlltkgagkgmvavfndmlrnladgnvfdssaogfatgagtgtdaikkngvesindft 720

QY 749 SSPGETINYTSNDYTLMDKIALSNPNSDEADRIKDELQAQAVVMTSGVFPNGGGEEM 808  
 DB 721 ssgetinytshndyltwdkialsnpnsdeadrilmdelagavvmtsgvfpngggeom 780  
 QY 809 LRKXGNDNSNAGDAVNEPDMKRAQYPOVFYNTYSGILHRLDHPFRMTTANETNSHL 868  
 DB 781 lrkxgndnsyagdaavnefpdmkraypovfyntysgilhrlthpfrmttanetnshl 840  
 QY 869 QPLNSPENTVAAYELTDHVNKDKMGNITVYVNPNTVATINLPSGKMAINATSGKVGSTL 928  
 DB 841 qplnspentvayeltdhvnkdkgmnlvynpnktvatlnlpsgkvalnatsgkvgastl 900  
 QY 929 GOAEGSVQVPEGISMMILHQEYSPDHGKK 956  
 DB 901 gqaegsvqvpjismilhqevspdhgk 928  
 RESULT 4  
 ID Y78514 standard; protein; 915 AA.  
 XX  
 AC Y78514;  
 XX  
 DT 05-MAY-2000 (first entry)  
 XX  
 DE Pullulanase amino acid sequence.  
 XX  
 KW Pullulanase; debranching enzyme; amylopectin; thermostable; corn;  
 KW genetically engineered variant; wheat; potato; wheat; manioc; rice;  
 KM starch; sugar production; high fructose syrup; high maltose syrup;  
 KM maltodextrin.  
 XX  
 OS *Bacillus deramificans*.  
 XX  
 PN WC200001796-A2.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PE 02-JUL-1999; 99MO-DK00381.  
 XX  
 PR 02-JUL-1998; 98DK-0000868.  
 XX  
 PA (NOVO ) NOVO-NORDISK AS.  
 XX  
 PI Bisgard-Frantzen H, Svendsen A;  
 XX  
 DR WPI; 2000-160767/14.  
 XX  
 PT Variant bacterial pullulanases and isoamylases having, e.g. increased  
 PT thermostability, used for converting starch from potatoes into high  
 PT fructose syrup -  
 XX  
 PS Claim 29; Fig 3; 116pp; English.  
 XX  
 CC This sequence represents a pullulanase amino acid sequence. Pullulanases  
 CC are debranching enzymes which can attack amylopectin. This sequence  
 CC represents a parent enzyme from which genetically engineered variant  
 CC debranching enzymes can be created. The genetically engineered enzyme has  
 CC an improved thermostability at a pH in the range of 4-6 compared to the  
 CC parent enzyme. The modified enzyme may alternatively or also have an  
 CC increased activity towards amylopectin and/or glycogen compared to the  
 CC parent enzyme. The enzyme variants are useful for converting starch to one  
 CC or more sugars. This is useful when starches such as corn, potato, wheat,  
 CC manioc and rice starch are used as starting materials in commercial  
 CC large scale production of sugars, such as high fructose syrup, high  
 CC maltose syrup, maltodextrins, amylose, G4-G6 oligosaccharides and other  
 CC carbohydrate such as fat replacers. The thermostable debranching enzymes  
 CC of the invention make it possible to perform liquefaction and debranching  
 CC simultaneously before the saccharification step. By debranching using the  
 CC thermostable debranching enzyme variants during liquefaction together  
 CC with the action of an alpha-amylase, the formation of panose precursors  
 CC is reduced, thereby reducing the panose content in the final product and

increasing the overall saccharification yield.

XX Sequence 915 AA:

Query Match 89.28; Score 4516; DB 21; Length 915;  
Best Local Similarity 93.08; Pred. No. 6,3e-287;  
Matches 866; Conservative 10; Mismatches 35; Indels 20; Gaps 4;

```

OY 30 DGNFTTIVHFCAPADYQPMWLMWMPKDGGAERDENQPDSPGAVASADIPGSPVYG 89
   |||||||
Db 1 dgttttivyhfrpadgqpwslmmpkdggaeydfnfpadsfgavasadiipnpsqvg 60
   |||||||
OY 90 IIVRTQMTKQVSDADRYIDLSKGNVWLVEGNSQIFVNEKDAEDAKPAVSNAYLDASNQ 149
   |||||||
Db 61 IIVRTQMTKQVSDADRYIDLSKGNVWLVEGNSQIFVNEKDAEDAKPAVSNAYLDASNQ 120
   |||||||
OY 150 VLVKLSQPLTIGBCKSGFTVVDOTANKDIPYTSVKDASLGGDYAVAVLAGTQHTFGSGDW 209
   |||||||
Db 121 VLVKLSQPLTIGBCKSGFTVVDOTANKDIPYTSVKDASLGGDYAVAVLAGTQHTFGSGDW 180
   |||||||
OY 210 APDNHSTLTKKVTNNLYOFSGDLPEGNTQYKALNDSWNNSYSPDNINLTVPAG----GA 265
   |||||||
Db 181 apdnhscltkkvtnnlyofsgdlpegnnyqykvamshs---agpvtsvqavwppgtrpylga 237
   |||||||
OY 266 HVTFSYIPSTHAYVDTLNNPNADLOVESGVKTDLVTVTLGEDPDVSHTLSTQTDGYQAKQ 325
   |||||||
Db 238 --twdgqlvntfals-----esgvktldvtlvtlgedpvtshstlsiqtdgyakq 284
   |||||||
OY 326 VIPRNVNLSQYVYSGDDLGNVTYQKATTFKVMAPTSTOVNVLVYDSATGSKTVPMHA 385
   |||||||
Db 285 VIPRNVNLSQYVYSGDDLGNVTYQKATTFKVMAPTSTOVNVLVYDSATGSKTVPMHA 344
   |||||||
OY 386 SGHGVEATVQNLNEMTYMEVETGOGSTRTAVDPYATATAPNGRGMIVDLAKTDPAWM 445
   |||||||
Db 345 sghyvewatvqnlenymtymevctggsttravdpayataleapngtrgmivdlaktdpagw 404
   |||||||
OY 446 NSDKHLPKNIIDEVYIEMDVYRDESIDPNSGMKNKGYLALTEKGTGKPDNVKTDLSLK 505
   |||||||
Db 405 nsdkhlpkniidevylemdivrdeidpnsgmknkgylaltekgtkgpdnvktgidslk 464
   |||||||
OY 506 OLGITTHQMLMPFASNSVDEIDPPQDNNWGYDPRNDVPEGQYATNANGNARKKEKEWYL 565
   |||||||
Db 465 qlgtlhwqimpyfasnsvdeidcptqdmgydprnydpegqyatanangnarkkekenwl 524
   |||||||
OY 566 SLHREHIGVNMDDVYNNHTFATQISDFDKIVPEYYRTMIQVLIIPDOYLEMKLAEKPMV 625
   |||||||
Db 525 slhrehigvnmddvynhthfataqisdfdkivpeyyrtlddegnytngsgtgcneiaaerpmv 584
   |||||||
OY 626 QKFIIDSLKYNVNEYHIDGFRFDLMLLAKDPTMSKASELHAINFGIALYGEPTWGTSTA 685
   |||||||
Db 585 qkfildslkyvneyhidgefrfdlmlalqkdtmskaaselhaingpialygepwtgtsa 644
   |||||||
OY 686 LPDQQLLTKGAKQKGVAVFNDNLNRLADGNVFEDSSAOCFATGACGLDIAKNGVEGSLN 745
   |||||||
Db 645 lpdqqltkgqkggvavfndnlrlnaldgnvfidsaagfagagylidaikngvegsln 704
   |||||||
OY 746 DFTSSPGETINVTSHDVTLMKDTALSNPNDEADRIKMDLQAOVAVVTSOGVFPMQOG 805
   |||||||
Db 705 dftsspgctinytshdvtlmdktalnsnpndeadrikmdlqaoavvtsogvfpmqog 764
   |||||||
OY 806 EEMLRKXKGNDSYNAGAVNEFDMSRKAQYPDVFNYSGLIHLRLDHPAFKMTTANETIN 865
   |||||||
Db 765 eemlrkxkgndsynagavnefdmsrkaqypdvfnysgllhlldhpaftmttanein 824
   |||||||
OY 866 SHLOFLNPEPNTVAELLDHVNKKDKGNTIIVYNNKNTVATNINLSGKWAINTAGKVE 925
   |||||||
Db 825 shloflnpepntvaelldhvnkkdkgntiivynnkntvatninsgkwaintagkve 884
   |||||||
OY 926 STLGAEGSVQVPGISMMILHQEVSPDHGKK 956
   |||||||
Db 885 stlgaegsvqvgpjismmilhqevspdhgk 915

```

```

RESUL 5
Y31747
ID Y31747 standard; Protein; 830 AA.
XX
AC Y31747;
XX
DT 22-NOV-1999 (first entry)
XX
DE Bacillus deramificans N-truncated pullulanase.
XX
KW pullulanase; enzyme engineering; starch; saccharification;
  mutant.
XX
OS Bacillus deramificans.
XX
FH Key
FT Misc-difference 134
  Location/Qualifiers
  /note= "encoded by GCN"
FT Misc-difference 135
  /note= "encoded by NNA"
FT Misc-difference 591
  /note= "encoded by CAN"
FT Misc-difference 592
  /note= "encoded by GCN"
FT Misc-difference 783
  /note= "encoded by ANA"
FT
FT
FT
PN W0945124-A2.
XX
PD 10-SEP-1999.
XX
XX 03-MAR-1999; 99WO-US04627.
XX
PR 04-MAR-1998; 98US-0034630.
XX
PA (GENEV) GENENCOR INT INC.
PI Miller BS, Shetty JK;
XX
XX WPI: 1999-540851/45.
DR
XX
XX
XX Claim 31; Page -: 49pp; English.
PT
PT New modified pullulanase for saccharification of aqueous liquefied
  starch
XX
XX The present sequence represents an N-terminally truncated
  pullulanase of Bacillus deramificans strain T89.117D, comprising
  amino acids 99-928 of the native mature enzyme (see Y31746). The
  native mature pullulanase (EC-3.2.1.41) has a pI of 4.1-4.5, a
  half-life of about 55 hr at 60 deg C and is capable of catalysing
  the hydrolysis of alpha-1,6-glucosidic bonds in amylopectin and
  pullulan. The invention relates to the discovery that B.
  deramificans pullulanase recombinantly produced in a Bacillus host
  is modified yet retains the ability to catalyse the hydrolysis of
  an alpha-1,6-glucosidic bond. The modification of the recombinant
  pullulanase appears to be a result of misprocessing of the signal
  sequence by a signal peptidase as well as susceptibility to
  extracellular proteolytic processing. The invention provides
  modified pullulanases, especially N-terminally truncated B.
  deramificans modified pullulanase, nucleic acids encoding them,
  vectors, host cells and methods for recombinant production. The
  modified pullulanases are useful in the starch industry for the
  saccharification of aqueous liquefied starch, forming syrups of
  high glucose and maltose content. Other uses (not claimed) are: as
  anti-staling additives in bread; in brewing low calorie beers;
  production of low calorie foods (where amylose is used as substitute
  for fat); to clarify fruit juices; in the preparation of
  oligosaccharides from amylopectin or of tetraoligosides from maltose;
  to condense mono- or oligosaccharides, or in the pharmaceutical
  industry (no details). When used for starch saccharification, the
  modified pullulanases reduce the amounts of by-products, especially

```



CC losses associated with reversion reaction products. When used with  
CC glucomylase, they make possible a higher level of dissolved solids  
CC in the material being treated, resulting in increased productivity  
CC and a more energy-efficient process.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from the *B. deramificans* pullulanase sequence given in  
CC figure 1.

SQ Sequence 826 AA;

Query Match 85.4%; Score 4323.5; DB 20; Length 826;  
Best Local Similarity 99.9%; Pred. No. 5.6e-284;  
Matches 825; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

OY 132 EDAAKPAVSNAIYLDASNOVLKLSOPLTLGEGXSGFTVHDDTANKDIEVTSVKDASLGOD 191
    |||||||
Db 1 edaakpavsnayldasnqvlklsqpltlgegxsqftvhddtankediprsvkdaslgqd 60

OY 192 VTAVALACTFQHI FEGSDMAPDNHSTLKKVTNNLYQFSGDLPEGNYQYKVALINDSMNN-S 250
    |||||||
Db 61 vtavalagfghlfggsdwapdnhscllkkvtcnlyqfsgdlpegnyqykvalindsmnps 120

OY 251 YPSDNINLTIVAGAGAHVFTSTIPSTHAYDTIINPNADLQVESGVTDLVYVTTGDEPDY 310
    |||||||
Db 121 ypsdninltivpagahvftstipsthaydtinpnadlqvesgvtclvltclgedpdy 180

OY 311 SHTSICITDGOAKOVIRRNRLNSQYYSGGDLGNTYOKATTFKWAPSTOVNVLVLY 370
    |||||||
Db 181 shlsicitdgoakqvrrnrlnsqysggdlgntcyqkattfkvwapstovnvilly 240

OY 371 DSATGSVTRKIYPMTAGSGHVEATVNONLENNYWEYVTGOGSTRTAVDPYATAIAPNGT 430
    |||||||
Db 241 dsatgsvtrkiypmtagsgvweatlvnqnlennwywevtgogstrtavdpatalapngt 300

OY 431 RGMIVDLAKTPAGANSKSHITPKNIEDVYIEMDVRFPSIDPNNGKMKKTYLLATEKG 490
    |||||||
Db 301 rgmivdlaktipaganskhiltpkniedevyemdvrfpsidpnngmkkkyallatekg 360

OY 491 TKGPDNVATGIDSLKQLGITHTVQLMPFASNSVDPTDPTDNMGVDPNRYDPEGOYATN 550
    |||||||
Db 361 tkgpdnvatgidslkqgithtvqimprfasnsvdetpdtqdwgydpnrydpegoyatr 420

OY 551 ANGNARIKEPKEMVLSLREHIGVNMDEVYNNHTFATQISDEDKIYEVYRTMIQVITPT 610
    |||||||
Db 421 angnarikefkemvlslhrehigvnmdevyynnhtfatqisdeldkiyeyyrtmiqvilpt 480

OY 611 DOVLEMKIAXRRPMYQKRTIISLKYWWEYHIDGRFDLMALGDTMSKASELHAINP 670
    |||||||
Db 481 dovlemkixaeprmyqkrtiislkywweyhidgrfdlmalgdtmskaaselhainp 540

OY 671 GIALYGEPMGTGTSALPDQDLTKGAKGMGVAVFNDMLRNALDGNVDSASQGFATGAT 730
    |||||||
Db 541 gialygepmgtgtsalpdqdltkgaqgmgyavfndmlrnaldgndvdsasqgfatagat 600

OY 731 GLTDAIKNGVEGINDFTSSPGETINVTSHDNTYLMDKIALSNPDSBADRIKMDQLAQ 790
    |||||||
Db 601 gltdaikngvegsindftsspgetinvtshdntylmdkialsnpdsbadrikmdelaq 660

OY 791 AVVMTSQGVPMQGEEMLRKXGNDNSYNAGDAVNEFDMSKAQYPOVFNTYSGILHNR 850
    |||||||
Db 661 avvmtsqgvpmqgeemlrkxgndnsynagdavnefdmskagypdvfnysgilnhr 720

OY 851 LDHPAFRRTANEINSHQFLNPSNPENTVAAYELTDHVNKDKMGNIIIVVNPNTVATINLP 910
    |||||||
Db 721 ldhpafrtaneinshqflnpsnpentvayaeltdhvnkdkmgniiivvnpnktvatlinlp 780

OY 911 SGKMAINATSGKVGESTLGQAEGSVQVPCISAMILLHOEVSPPDHGRK 956
    |||||||
Db 781 sgkmainatsgkvgestlgqaegsvqvpqisammillhgevsppdhgrrk 826

```

RESULT 7

```

W37372
ID W37372 standard; Protein; 922 AA.
XX
AC W37372:
XX
DT 11-MAR-1998 (first entry)
XX
DE Trimmed enzyme protein.
XX
KW Bacillus sp. APC-9603; trimmed enzyme; substrate specificity;
   heat resistance; acid region; production.
XX
OS Bacillus sp.
   JP09271385-A.
XX
PD 21-OCT-1997.
XX
PF 01-APR-1996; 96JP-0134492.
XX
PR 01-APR-1996; 96JP-0134492.
XX
PA (AMANO ) AMANO PHARM KK.
DR WP1: 1998-003022/01.
DR N-PSDB: T97233.
XX
PT DNA encoding a trimmed enzyme - used in the production of a trimmed
   enzyme with wide substrate specificity
XX
PS Claim 2; Page 9-10; 15pp; Japanese.
XX
CC The present sequence represents a trimmed enzyme isolated from
   Bacillus sp. APC-9603. The present specification also describes:
   (1) a vector containing the DNA encoding a trimmed enzyme; (2) a
   recombinant microbe transformed by the DNA; and (3) preparation of
   a trimmed enzyme in which the microbe of (2) is cultured and the
   CC trimmed enzyme is collected from the culture. The method can prepare
   CC a trimmed enzyme of wide substrate specificity having heat resistance
   CC and being active in acid region efficiently.
SQ Sequence 922 AA:

Query Match 82.0%; Score 4148; DB 19; Length 922;
Best Local Similarity 83.7%; Pred. No. 4.9e-272;
Matches 772; Conservative 63; Mismatches 85; Indels 2; Gaps 2;

OY 30 DGNFTTIIVHYFCPAGADYQPMSLMMWPKDGGAEYDFNQPADSFGAVASADIPGNSOYG 89
    |||||
Db 1 dgnttnvlihyfrpgdyqswslmmwpegdgnynfn-gtdsygelanvslpsspkvg 59

OY 90 IIVRTQDWKQVDSADRYTDLSSKNGEYWLVEGNSQIFYNKDKMEDAKKPAVSNAIYLDASQ 149
    |||||||
Db 60 iivrtqdwkqvsqdytldlskngewlvegnsqifynkdkmedakpavsnayldapnk 119

OY 150 VLVKLSQPLTLGEGXSGFTVHDDTANKDIPYTSVKDASIGODVYTAVALACTFQHI FEGSPW 209
    |||||||
Db 120 vlvklsqpltlgegxsqftvhddtankediprsvkdasigodvtavalactfghlfggsdw 179

OY 210 APDNHSTLKKVTNNLYQFSGDLPEGNYQYKVALINDSMNNSYPSDNIINLTIVAGAGAHVTF 269
    |||||||
Db 180 apdnhscllkkvtcnlyqfsgdlpegnyqykvalindsmnnsypsdniinltivpagahvtf 239

OY 270 SYIPSTHAYDTIINPNADLQVE-SGVRKDYTVTLTGDEPDVSHLSTIOTDGOAKQVIRP 328
    |||||||
Db 240 syipsthaydtinpnadlqvsgvrtkdytvtltgdepdvshlsitotdgoakqvirp 299

OY 329 RNLVNSQYYSGGDLGNTYOKATTFKWAPSTOVNVLVLDASATGSVTKIYPMTAGSGH 388
    |||||||
Db 300 rnlvnsqysggdlgntcyqkattfkwapstovnvllvldasatgsvtkiypmtagsgh 359

OY 389 GWMEATVNONLENNYWEYVTGOGSTRTAVDPYATAIAPNGTRGMIVDLAKTDPAGWNSD 448

```

D	b		360	gwwavetvpqnlennwylyevtcggsctrcvavdpayatalapngtcrgmvdlacltoprtcksd	419
O	y		449	khiPRKNIkEDEvTYEMQVDRDSIDPNSGMRKKKRYLALTEKGTCGRPNWKTGIDSLQLQG	508
D	b		420	 khmrPRkIedeyIYemhvrdtsIdsgnmrltkyxlalcelkytlcrgpanvktyvdsIkqIg	479
O	y		509	ITHVOLMPEFVASNSVDETPDPTDMMKGYDPRRNYDVPEGQYATINANGNARIKEFEKEMVLSLH	568
D	b		480	lchvgylqrvaafnsavdecdprcywgydprnrynpesgyatcdangctrickefkemvlslh	539
O	y		569	REHGIVNNDDVVYNNHTFAQTQISDFPKIYPEYYIRTMIOVIILPTDOVLDMKLKAERPMYQKR	628
D	b		540	rnhlgvmmdvvynhlftelqIsdlfdlkIypqyyrlddagnylnsgtgnevaaerpmvqkf	539
O	y		629	IIDSfKYVNEVEYHIIDGFEDLMALIGDKTMSKAASELHAIPGIALGEPWTGGTSALPD	688
D	b		600	lIdstkyvveeyhlidgfifdtmlllgkdImakaqeqlhaladpyalaIYsepvcqtaselpc	639
O	y		689	DQLLTGCAQCKMGVAVFENDLRNALDGNVFDSSAQGFATGATGLTDIAIKNGVEGSINDFT	748
D	b		660	dqllltkygqkgmyavfnndlrngldgnvfdaesqfatgatgtldvIkkyvegslndfc	719
O	y		749	SSPGETINYVSHNNTYLMDKIALSNPNDSPADIKKDELAQAANVMYSOGVFPFGQGEEM	808
D	b		720	sppccelliyvshnhytlwdIiaqsnprdsaadIkmdelaqavvnvsqvrplmggeem	779
O	y		809	LRLXKCNDNSYNAGDAVNEPFMSRKAQYPDYFNYSGLHLRLDHPAFRMTYANETINSHL	868
D	b		780	lrlcggnasnaynaegdavefdawfrkaqsydvfnysglhlrlclhpfrfmlntuanqlkehI	839
O	y		869	QELNSPEVTAVYELDLHVNRKKMKGNIIIVVNPKRTVATINLPSSKMAIMNTSGKVGESTL	928
D	b		840	qtIdspndtveyeltnhnakkkwgnlvivlynpxtaetvnlpsqkwaInatngktgestl	899
O	y		929	GQAEGSVOVPGISMMILHQEVS	950
D	b		900	shaeghvqvpglsmmlhqetn	921
R	E	S			
I	D		Y78513	standard; protein; 862 AA.	
A	C		Y78513;		
D	T		05-MAY-2000	(first entry)	
X	D			Pullulanase (pulB) amino acid sequence.	
X	D				
X	S		Bacillus acidopullulyticus.		
X	O		MO200001796-A2.		
X	P		13-JAN-2000.		
X	F		02-JUL-1999;	99WO-DK00381.	
X	P		02-JUL-1998;	98DK-0000868.	
X	P		(NOVO ) NOVO-NORDISK AS.		
X	P		Bisgaard-Frantzen H., Svendsen A:		
X	P		WPI; 2000-160767/14.		
D	R		N-PDB; 289960.		
X	X			Variant bacterial pullulanases and isoamylases having, e.g. increased	

[illegible]

```

Db      511 gnytngsvgnelaterpmvqkfvldsvkwyvkeyhldgfrtdlmaligkctmakiskel 570
QY      666 HAINFGIALYGEPMWTGTSALPDQLTRKGAOKGMGAFAFNDRNALDGNFESSAQGF 725
Db      571 haiopgviylgepwtgsgsgslsdqilvckgqkylgivyfndnirngldgnvfdksagqf 630
QY      726 ATGATGTLDAIKNGVESSINDFTSSPGETTINTVTSHDNVTYLMKIALSNPNSEADRIKM 785
Db      631 atgdpgnvnvlkxrwmgslsdfstapselinytshdmcltcklisaenpdtqadrlkm 690
QY      786 DELAGAVVMTSGVPEMGCGEMLRXKGNDSYNAGDAVNFEDMSRKQAYDVFNYYSG 845
Db      691 delagavvfcsqvpfmggsemlrtkgnndsynagdsnvqfidsrkagfenvtdyysw 750
QY      846 LIHLIDHPAFRMTTANESHLQFLNSPENTVAYELTDHVKDKMGNIIVYVNPRTVA 905
Db      751 lhlrdhnpafmttadqikgnltfidsptntvatefelnahndkwnknlivmynpntkag 810
QY      906 TITLPSGKMAINATSGKVGESTLGOAEGSVQVPGISMMILHDESPD 952
Db      811 tltpsgmwtlvglygnvgkslghvngltvevpaltstllhbgtsed 857

```

## RESULT 9

ID W49871 standard; Protein; 772 AA.

AC W49871;

DT 21-DEC-1998 (first entry)

DE Thermotoga maritima 6Gp3 pullulanase.

KW glycosidase; 6Gp3; thermostable enzyme; oligosaccharide; glucose;

XX sugar; baking; textile; detergent; pullulanase.

OS Thermotoga maritima strain 6Gp3.

XX MO982479-A1.

PD 11-JUN-1998.

PF 08-DEC-1997; 97WO-US22623.

XX 10-OCT-1997; 97US-0949026.

PR 06-DEC-1996; 96US-0056916.

XX (DIVE-) DIVERSA CORP.

PA Bylina EJ, Lam DE, Mathur EJ, Swanson RV;

DR WPI: 1998-362407/31.

DR N-PSDB: V36920.

XX Glycosidase enzymes from organisms of the genera *Staphylothermus*,

PT *Pyrococcus* and *Thermococcus* - for deriving sugar from

PT oligosaccharides; useful in the e.g. food processing, textile or

PT baking industries

XX Claim 1; Fig 14a-e; 92pp; English.

XX This is the amino acid sequence of pullulanase 6Gp3, deduced from

CC an isolated polynucleotide (see V36920) of a *Thermotoga maritima*

CC clone (6Gp3) that grows optimally at 85 degC. The sequence shows

CC 72% amino acid identity to an alpha dextran 6 glucanohydrolase of

CC *Caldocellum saccharolyticum*. The invention provides 18

CC polynucleotides (see V36907-24) coding for thermostable glycosidases

CC (see W49871-75) having glucosidase, alpha-galactosidase,

CC beta-galactosidase, beta-mannosidase, beta-mannanase, endoglucanase

CC or pullulanase activity. Vectors and host cells are also claimed.

CC A method is provided for producing the enzymes by recombinant

CC techniques. A claimed method for generating glucose from soluble

CC cell oligosaccharides comprises contacting a sample (selected from

CC dairy products, fruit juice, detergent, textile, guar gum, animal

CC feed, plant biomass or waste product) containing oligosaccharides

CC (selected from maltose, cellobiose, lactose, sucrose, raffinose,

CC starchose, verbascose, cellulose, starch, amylose, glycogen,

CC disaccharides, polysaccharides and pullulan) with one of the

CC claimed glycosidases such that glucose is produced.

XX

SQ Sequence 772 AA:

Query Match 29.0%; Score 1465; DB 19; Length 772;

Best Local Similarity 36.3%; Pred. No. 1e-90;

Matches 326; Conservative 138; Mismatches 270; Indels 164; Gaps 20;

QY 86 SOVGIVRTQDM--TRKVSADRYIDLSKGN--EWMALVEGNSOIFNEKDAE-----DAAP 137

Db 4 tkvgilvrlnewqakavakvdrleldgkaevallgqveellfeyekpdsprlffagarsn 63

QY 138 AVSNAYL-----DASNOVLVKLSQLPLTGKXSGGTFVHDPTANKDIPVTSKDAASLGQDVT 193

Db 64 kvleatfltpvtdkkkelkvt-----vdgkelpvsrveka-----dpt 102

QY 194 AVLAGTFQHI FGGSQDAPDNHSTLLKKTNNLTYQESGDLPEGNYQYKVALNDSWNNSTYPS 253

Db 103 -----divtlny-vrlvlsesl----- 118

QY 254 DNINLTPVAGAHVTFSTYIPSTHAVYDTINNPNADQVSGVGTDLVYTLGSDPDVSHRT 313

Db 119 -----keeldlrkd 126

QY 314 LSIQTDGYQAKOYIPRNVLNSSQYYSGGDLGNWTYOKATTFKVNAPTSTOVNVLLYDSA 373

Db 127 velliegykparvclmelld--dyrydg-elgavyspeklffvsvpskvvllfkng 183

QY 374 TGSVY- KIVPMPTASGHWEDATVYVNONLEWNYWYEVTTGGGSTRFVADPYATAIAPRTNG 432

Db 184 edtepygvvnmeykgnvveavvegldgvyfylyglenygkirtvdpskavaynnges 243

QY 433 MIVDLAKTDPKMSNDKHTPKNIEDEVITYEMDVRPESIDPNSGMMKCKYALATETKGRK 492

Db 244 avvnlarlnpegwendrpkiegyedaillvelhadielglsnsvknkylglileenk 303

QY 493 GPDNVKTGIDSLKOLGITVVOILMPVFNASNSVDETPDTPD---NMGXDPNRYDVPECOVAT 549

Db 304 gpgvgtvtglshlvelgvtvnhlplfdlytgedldkfykynwsgdpylflwpegryst 363

QY 550 N-ANGNARIKERKENVLSLHREHIGVNMVYVNHFEAT-QISDPDKIVPEYRYRTMIOVI 607

Db 364 dphnphtrirevkenmkalkhnglgyimdmvfphtyglgelsafdqtpyyfyr----- 418

QY 608 IPTDQYLEMK-----LXAREPMVOKFIIDSLKRWVMEYHIDGRFPLMLLGDQTSKKA 662

Db 419 dktgaylnsgcgnvlassepmmrkfivdvtlywvkeyhldgfrtdlmaligkctmakiskel 478

QY 663 SELHAINPGIALYGEPMWTGTSALPDQLTRKGAOKGMGAFAFNDRNALDGNFESSAQGF 722

Db 479 ralhkidprlllygepwwgwapl-----fpgksdvglthvaeinfdefridaigsvtnpsv 534

QY 723 OGFATGATGLTDAIKNGVECSIN-----DFTSSPGETTINTVTSHDNVTYLMKIALSNP 775

Db 535 kgfmggygkckekikrgvgsinydgkllksfaldpheetinyachdnltlwknylaak 594

QY 776 ND-----SEADRIKMDLEAOAVVMTSGVPEMGCGEMLRXKGNDSYNAGDAVNFEDMSRKQAYDVFNYYSG 845

Db 595 adkkkewteelknagklsagallitsgvpflfnggdqfcttnfnfnasnapistngif 654

QY 831 SRKAQYPDVFNYYSGLIHLRDLHPAFRMTTANESHLQFLNSPENTVAYELTDHVKDKMGNIIVYVNPRTVA 890

Db 655 etklgldvrfnyhkgkllkrlkehpaftrlknaeikkhlhlpqgrvrvfamlkdhagqpb 714

QY 891 WGNIIIVYVNPRTVATINLPSGKMAINATSGKVGESTLGOAEGSVQVPGISMMILHDESPD 948









```

RESULT 14
W73553 standard; protein; 1938 AA.
XX
AC W73553;
XX
DT 08-MAR-1999 (first entry)
XX
DE Full length pullulanase protein.
XX
KW Pullulanase; variant; bleach-containing detergent; oxidation resistant.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT MISC-difference 121..137
FT /note= "nucleotides encoding these residues are not
FT present in the coding sequence given in the
FT specification"
FT MISC-difference 1050..1065
FT /note= "nucleotides encoding these residues are not
FT present in the coding sequence given in the
FT specification"
XX
PN JP10327866-A.
XX
PD 15-DEC-1998.
XX
PE 30-MAY-1997; 97JP-0141596.
XX
PR 30-MAY-1997; 97JP-0141596.
XX
PA (KAOS ) KAO CORP.
XX
DR WPI: 1999-099031/09.
DR N-PSDB: V62881.
XX
FT New mutant pullulanase - useful in bleach-containing detergents
XX
PS Claim 13; Page 10-17; 19pp; Japanese.
XX
CC This sequence represents the Bacillus pullulanase protein. The invention
CC relates to a mutant pullulanase in which at least one methionine residue
CC in the pullulanase is deleted or replaced by an optional amino acid
CC residue other than Cys and Met and/or at least one alanine residue in the
CC amino acid residues constituting pullulanase is deleted or replaced by an
CC optional amino acid residue other than Ala. The mutant pullulanase is
CC useful in bleach-containing detergents. The pullulanase is highly
CC resistant to oxidation.
XX
SQ Sequence 1938 AA;

```

Query Match 16.2%; Score 817.5; DB 20; Length 1938;  
 Best Local Similarity 25.8%; Pred. No. 2,5e+46;  
 Matches 284; Conservative 153; Mismatches 369; Indels 293; Gaps 50;

```

QY 35 TIVIVFCPCADYOPMSLMMW-----PKD---GGCAEYDFNPADSFCAVASADIGNPS 86
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 835 tlrhlyqtrtdnsyenlgilwgdvaarsenwpsgqtrfpaqnvtd-ygaayvvelaadaq 893
QY 87 OVG-IIVRTQMTKDVASDRYIDLKSG--NEVWLVEGNSQIF----- 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 894 nlgflvltntgdkd-ggdkavelfspdneilwksdevflpeyvdipantvrihyer 952
QY 126 -----VYEKDAE-----DAKPAVSNAFLDLSNOVLVLSQPLTIGEXX 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 953 tnaadyegwglwmedvepsdgwpngaadaagigkygayd-----ixlk----- 997
QY 165 SGFTVHDDTANKDIPVTSVKDASLCO--DVTAVLAGTFQHF--GSGDPAAPDNHSTLKK 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 998 -----edaak-igflfvnksgsgqtdmtdmlkqynqdlvkeged-----k 1038

```

```

QY 221 VTNNTLYQFSGDLPEGNYYQKVALNDNSMNSYPSDINILTV-----PA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1039 vytiny-----gtvplalvsgev-lsckl-----slftftrtegidleelkeqlckav 1086
QY 263 GGAHYTFSTIPSTHAYVDITINNPNADLOYESGVKT-----DL-----VTYVLGEPPDVS 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1087 dgndvsft-----dvtle-geklvthgefdleklpfsvty-----ld 1123
QY 312 HTLSIOTDGYOKOVIPRNVLNSQYYVSGDDIGNTYTOKATFTKVMAPTSTQVNVLYLD 371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1124 rltsvks-gwk-----ldemyaydgkigaehgedgtalckvpskadvsvvlyld 1173
QY 372 SA-TGSVTKIVPMTASGHCWMEATVNO-----NLEMYTMEVTSQSTRAVDYATA 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1174 kvdqnevdtliemvkgdvswsklktkntgldslkgyyhyethgtvntaldpyaks 1233
QY 425 IAP-NETRG-----MIVDLAKTDPAQWMSDKHITP--KNIEBEVYENDVYRFSIDPMS 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1234 maawneaadkvkgaalvlgslg--eladyadipgfeekredllyevhvridfstdpnl 1290
QY 476 GKMKN---GKYALATTEKGTGKGPDNVKTGIDSLKOLGITHVOLMPV---FASNSVDE---- 525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1291 gedlkagfifsfvek-----ldyqelgvthqllpwmssyfsnefesgem 1339
QY 526 ---IDPTQDNMGYDERNVDPREGQYATN-ANGNABIKFEKENVLSLHREHIGVNDVYV 580
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1340 levasctgynywgydphnyfslsgmysenpedelrkefknlneihkrmdgvidvtf 1399
QY 581 NHTFATQISDFPKIPEYRYRTMIOVIIPDOVLEKKLAERMVOKFIIDSLAKYVNEY 640
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1400 nhl--aqvnlfedlvpnyyh-fmdadgtrptsfggrrlgtlhemssrvlvdsikhwvdey 1456
QY 641 HIDGFRFDLMALLGKDTMKAASELHAIPGIALYGEPT--GGSALPDQDLTKGAOK 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1457 kvdgrfrdmagddaesiglafeakklpnulymigegwvtfsgdcegygaadqymgy 1516
QY 699 GMGCAVFNDNLRNALDGNVFDSSAGFATGAGLTDAIKNGVGSINDP--TSSPETINTY 757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1517 teaagsfsdfeirnelksgfsgseqprflicgavnvqgldonlkagphfmdadpgdvvqy 1576
QY 758 VTSHDVNTLMDKIFALSNPNDSADRIKMD-----ELAQAVVMVSGQVPPMGGEELRRK 812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1577 leahndlltydvaagikkdpelaendlelhrk-irgnamvllsglcalihagqefgrck 1636
QY 813 -----GDN-----DNSYNAGDAVNEFDMSRKA---OYPPVFN- 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1637 qwrpaateapykstymtdadgnpfyvyfihdsydsdlinrtfdekatdaekyp-vmyv 1695
QY 842 ---YSGLIHLRLDHPAFRMTTANEINSHLOFLNSPE-----NTVAYELTDH 885
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1696 trdytaglilelrssdafrlgsrelvdsvntwmdapeikeqdlvayrsvstagsyelt 1755
QY 886 VN-----KDKMGNIIVYVNNKNKYATINLPSG-----KKAINATSKGYE 925
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1756 vnaatsrtlltqgdlllegvvvvaadeaanvagaapagfeltaegiltleptltvvrvge 1815
QY 926 ST-----LGOAGESVQVPG 939
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1816 qegtdpgdgddgndtpppg 1834

```

```

RESULT 15
W09255
ID W09255 standard; Protein; 1938 AA.
XX
AC W09255;
XX
DT 27-MAR-1997 (first entry)
XX
DE Bacillus alkaline amylopullulanase.
XX
KW Amylopullulanase; alpha-amyase; pullulanase; detergent additive.
XX

```

CC	anylipoululanase, alpha-amylase and pullulanase can each be
CC	expressed in transformed host cells using vectors contg. the
CC	appropriate gene fragments. The enzymes are useful as additives
CC	in dish-washing and clothes detergent compans.
CC	
xx	
Sequence	1938 AA:
Query Match	16.1% Score 816.5; DB 17; Length 1938;
Best Local Similarity	25.8% Pred. No. 2.9e-46;
Matches 284; Conservative 153; Mismatches 369; Indels 293; Gaps 50;	
QY	35 TIIVHYRCPACDYOPLMSLMW-----PKD---GGCAEYDFNOPADSFCAVASADIPGNS 86
DB	835 LIRIHGFRLDSYENIGLWIGDVAAPENWPSGRLPQAGNVLD-YGAYDVDEIADAQ 893
QY	87 QVG-IIRTOQWTRDVASDRYIDLSKG--NEWLVEGNSQIF----- 125
DB	894 NIGLIVLNTLTGDDK-GQDKAVELFSPDLNGLWIKGSDEVFLYEPVDIPANTVRIYER 952
QY	126 -----YNEKDAE-----DAKPNVSNVYLDASNOVLKLSQPLTLEGX 164
DB	953 LNADEYEGKLWNEWDEEVPADGYPGNGAADAAGIKGYGEAYD-----IKL- 997
QY	165 SGPTVHDDTANKDIPVTSVKDASLQ--DVTAVLAGTFQHLF--GGSMDAPDNHSTLLK 220
DB	998 -----EDANK-IGLFLVnkqsggqldgmfdmklkynqlfvkkged-----k 1038
QY	221 VTNNLYQFSGLPEGNQYKALNDSMNNSYSPSDNINTV-----PA 262
DB	1039 VYTPY---GLVPLALVSQEV-LSDKL-----ISLFTREGLDLEELKEGLKDV 1086
QY	263 GGAHVFESYIPSTHAAYVDITNNPNADLOVEESGVT-----DL---VTTLGEDPDVS 311
DB	1087 dgnvstf-----dvrlr-geklvnhgfdleklpfsvty-----ld 1123
QY	312 HTLSIQFDGQAKOAVIPRNVLNSQVYYSGDGLGNFYTKQATTKKVNAPSTOVNVLVD 371
DB	1124 RLIVSKS-GVK-----LIDEMAYDGLIGLELHEDGATCLIKVSPKADNSVVLVD 1173
QY	372 SA-TGSVTKIAPMTASGHWGEATVNO-----NLENNYVYEVTVGGOSTRTAVDPATA 424
DB	1174 kvdgnevvdltiemkkgdrtgvsvkrlcdnlgldsklyyylyvalthdvtlraldpyaks 1233
QY	425 IAP-NGTRG-----MIVDLAKTDPACMNSDKHITP--KNIEDEYIEMOVRDPSIDPS 475
DB	1234 maanneagdkvgkaalrvdsgisyp---eldyadltpgfekedcllyevhvrldtsdphl 1290
QY	476 GMRKK---GKYLALTEKGTGPDNVKGTGLOSILKOLGITVHOLMV---FASNSVDE--- 525
DB	1291 gedlkaqfgtfafevex-----ldyIgeIgyrlthqlIprmsyfyaneIegeerm 1339
QY	526 -----TDPTQDNMGVDPENRYDPEGOYATN-ANGNAKRIKEPEEMVLSLHREIIGNMDVY 580
DB	1340 lqvastqtnynwgdyphnyfislsgmysempedbelrIkefnllneIhkrctmgvldvlf 1399
QY	581 NHFPATQISDFDKIVPEYRYRTMIOVLIIPQVLEMLKXERPAQFIIDSLYVNEY 640
DB	1400 nhtc-aqvhlfedvlpnyyh-fmdadgtprltsfyggfclgtchemsrrvlyvslkhwvdey 1456
QY	641 HIDGFRFDLALLGKDTMSKAASELHAINPICALYGEPMW--GCTSLPDDQLTLTKGAK 698
DB	1457 kvdgrfrdmngdhdaeslqldaekklpnlvmlgsqvwclfgadegepvqaadqgmymqy 1516
QY	699 GKGVAYVNDLRLNALDSGNVDSAGQATGATGTLDAIKNGVGSINDF-TSSGEGTINT 757
DB	1517 teaqgsfsdfeirnelkqsgsgqgrflfitygaavnvqIfdnIkaqphfmadqpgdvvyq 1576
QY	758 VTSIHDTLLMDKLTALSNPNPNSSEADRIKMD-----ELAQAVYMTSQGVPPMGGEMLRX 812
DB	1577 IeahdmlllydvlaqsglkkkpeelaendleIhkrItyvgnamllysqglaflhaqgefgtck 1636
QY	813 -----CGN-----DNSYNADDAVNEFPWSRKA---QYPDVFN- 841





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 07:58:36 ; Search time 51.08 Seconds

(without alignments)  
359.545 Million cell updates/sec

Title: US-09-262-126c-2

Perfect score: 5060  
Sequence: 1 MAKKLYCLVCLVLTMAF.....VPGISMMILHQVSPDHGKR 956

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued\_Patents\_AA.\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the total being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	4888.5	96.6	928	1	US-08-474-140-11 Sequence 11, Appl
2	4888.5	96.6	928	1	US-08-477-630-11 Sequence 11, Appl
3	4888.5	96.6	928	1	US-08-472-293-11 Sequence 11, Appl
4	4888.5	96.6	928	1	US-08-474-545-11 Sequence 11, Appl
5	4888.5	96.6	928	2	US-08-478-341-11 Sequence 11, Appl
6	4888.5	96.6	928	3	US-08-996-733-11 Sequence 11, Appl
7	4888.5	96.6	928	4	US-08-860-339-18 Sequence 18, Appl
8	543	10.7	618	4	US-08-961-083-72 Sequence 72, Appl
9	509	10.1	746	1	US-08-476-519-11 Sequence 11, Appl
10	509	10.1	746	5	PCR-US95-09333-11 Sequence 11, Appl
11	509	10.1	777	1	US-08-476-519-2 Sequence 2, Appl
12	509	10.1	777	5	PCR-US95-09333-2 Sequence 2, Appl
13	487	9.6	751	6	5457037-5 Patent No. 5457037
14	482	9.5	750	6	5457037-3 Patent No. 5457037
15	475	9.4	726	3	US-09-129-075-4 Sequence 4, Appl
16	467.5	9.2	740	2	US-08-410-784A-2 Sequence 4, Appl
17	465.5	9.2	740	2	US-08-410-784A-2 Sequence 2, Appl
18	434	8.6	772	2	US-08-410-784A-5 Sequence 5, Appl
19	248	4.9	556	2	US-08-505-377-1 Sequence 1, Appl
20	248	4.9	556	3	US-08-798-269-1 Sequence 1, Appl
21	237.5	4.7	589	1	US-08-399-646-2 Sequence 2, Appl
22	237.5	4.7	589	1	US-08-607-321-2 Sequence 2, Appl
23	237.5	4.7	589	2	US-08-607-321-2 Sequence 2, Appl
24	237.5	4.7	589	2	US-08-605-501-2 Sequence 2, Appl
25	237.5	4.7	596	1	US-08-399-646-12 Sequence 12, Appl
26	237.5	4.7	596	1	US-08-607-321-12 Sequence 12, Appl
27	237.5	4.7	596	2	US-08-961-240-12 Sequence 12, Appl

28	237.5	4.7	596	2	US-08-605-501-12 Sequence 12, Appl
29	231	4.6	588	2	US-08-339-715A-5 Sequence 5, Appl
30	229	4.5	588	2	US-08-339-715A-1 Sequence 1, Appl
31	229	4.5	588	2	US-08-339-715A-3 Sequence 3, Appl
32	223.5	4.4	597	1	US-08-399-646-4 Sequence 4, Appl
33	223.5	4.4	597	1	US-08-607-321-4 Sequence 4, Appl
34	223.5	4.4	597	2	US-08-961-240-4 Sequence 4, Appl
35	223.5	4.4	597	2	US-08-605-501-4 Sequence 4, Appl
36	223.5	4.4	598	1	US-08-399-646-14 Sequence 14, Appl
37	223.5	4.4	598	1	US-08-607-321-14 Sequence 14, Appl
38	223.5	4.4	598	2	US-08-961-240-14 Sequence 14, Appl
39	223.5	4.4	598	2	US-08-605-501-14 Sequence 14, Appl
40	223	4.4	588	2	US-08-339-715A-4 Sequence 4, Appl
41	214	4.2	676	3	US-08-947-965-71 Sequence 71, Appl
42	205.5	4.1	600	1	US-08-374-155A-10 Sequence 10, Appl
43	205.5	4.1	600	2	US-08-785-396-10 Sequence 10, Appl
44	205	4.1	629	1	US-08-374-155A-4 Sequence 4, Appl
45	205	4.1	629	2	US-08-785-396-4 Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-474-140-11  
; Sequence 11, Application US/08474140  
; Patent No. 5721127  
; GENERAL INFORMATION:  
; APPLICANT: DEMEER, PHILIPPE  
; APPLICANT: AMORY, ANTOINE  
; TITLE OF INVENTION: PULLULANASE, MICROORGANISMS WHICH  
; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
; STREET: 2000 K Street, N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,140  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: William F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-41  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 429-0625  
; TELEFAX: (202) 293-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 928 amino acids  
; TYPE: amino acids  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-474-140-11

Query Match 96.6%; Score 4888.5; DB 1: Length 928;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 926; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
QY 30 DGVTTTIIYHFPADYQPMISLMMWPKDGGAEYFENPADSFGAVASADIRGNSQVG 89  
|||||

Db 1 DGNNTTIIIVHFCPCAGDYQWPWSLMMMPKDGGAETDFNOPADSGAASADIPGNPSQVG 60  
QY 90 IIVRTQDMTKVDSADRYIDLSKNGEVLVEGNSQIFYNKEKAEDAAAPASNAVLADSNQ 149  
Db 61 IIVRTQDMTKVDSADRYIDLSKNGEVLVEGNSQIFYNKEKAEDAAAPASNAVLADSNQ 120  
QY 150 VLVLKSPPLTIGEGXSGFTVHDDTANKNDIPYTSVKDASLGDVYAVLAGTQHIFFGSDW 209  
Db 121 VLVLKSPPLTIGEGXSGFTVHDDTANKNDIPYTSVKDASLGDVYAVLAGTQHIFFGSDW 180  
QY 210 APDNHSTLLKAVTNKLVQFSGDLPEGNQYQVVALNDSNN-SYPSDNIINTVPAGAHVT 268  
Db 181 APDNHSTLLKAVTNKLVQFSGDLPEGNQYQVVALNDSNN-SYPSDNIINTVPAGAHVT 240  
QY 269 FSYIPSTHAYVDTINNPADLOVESGKVTDLVTVTLGEDPVSHTLSIQTDGYAKQVITP 328  
Db 241 FSYIPSTHAYVDTINNPADLOVESGKVTDLVTVTLGEDPVSHTLSIQTDGYAKQVITP 300  
QY 329 RNVLNNSQYYSGGDLGNTYQKATTFKVMAPSTOVAVLLYDSATGSVTKIVPMTASGH 388  
Db 301 RNVLNNSQYYSGGDLGNTYQKATTFKVMAPSTOVAVLLYDSATGSVTKIVPMTASGH 360  
QY 389 GWEATVNONLENNYVYMEVYGCGSTRPAVDPAATAPNCTRCMIYDLAKTDPAGWNSD 448  
Db 361 GWEATVNONLENNYVYMEVYGCGSTRPAVDPAATAPNCTRCMIYDLAKTDPAGWNSD 420  
QY 449 KHITPKNIEDEVYIEMVDROFSIDPNSGMMKKGYLALTEKGTGPDVVKGTGIDSLQOLG 508  
Db 421 KHITPKNIEDEVYIEMVDROFSIDPNSGMMKKGYLALTEKGTGPDVVKGTGIDSLQOLG 480  
QY 509 ITHVOLMPVFASNSVDETPQDMNGGIDPRNYDVPBGQYATNANGNARIKFEKEMVLSLH 568  
Db 481 ITHVOLMPVFASNSVDETPQDMNGGIDPRNYDVPBGQYATNANGNARIKFEKEMVLSLH 540  
QY 569 REHIGVNMDDVYNNHTFATQISDPDKIVPEYVYRTMIOVLIPTDQVLEMKLAERPMYQKF 628  
Db 541 REHIGVNMDDVYNNHTFATQISDPDKIVPEYVYRTMIOVLIPTDQVLEMKLAERPMYQKF 600  
QY 629 IIDSILKAVNNEYHIDGRFIDLALGKDTMSKAASELHAIPGALIGEPMVTGTSALPD 688  
Db 601 IIDSILKAVNNEYHIDGRFIDLALGKDTMSKAASELHAIPGALIGEPMVTGTSALPD 660  
QY 689 DQLLTKGAQKMGVAVFENDLRNALDGNVFPDSSAOGFATGTLTDIKNVEGSINDFT 748  
Db 661 DQLLTKGAQKMGVAVFENDLRNALDGNVFPDSSAOGFATGTLTDIKNVEGSINDFT 720  
QY 749 SSPGETIYVTSHDNYTLMDKIALSNPNDSEADRIKMDELAQAAYMTSQGVPFMOGGEBM 808  
Db 721 SSPGETIYVTSHDNYTLMDKIALSNPNDSEADRIKMDELAQAAYMTSQGVPFMOGGEBM 780  
QY 809 LRXXGNDNSYNACDVAWEFPMKRKAQYPDVFNYSGLHLRLDHPAFRMTTANENSHL 868  
Db 781 LRXXGNDNSYNACDVAWEFPMKRKAQYPDVFNYSGLHLRLDHPAFRMTTANENSHL 840  
QY 869 QFLNSEPNTVAYELTDHYNKDKMGNIIIVYNPNKTVAATINLPSGKMAINATSGVGESETL 928  
Db 841 QFLNSEPNTVAYELTDHYNKDKMGNIIIVYNPNKTVAATINLPSGKMAINATSGVGESETL 900  
QY 929 GOAEGSVQVPGISMMIILHOEVSPPDHCK 956  
Db 901 GOAEGSVQVPGISMMIILHOEVSPPDHCK 928

RESULT 2  
US-08-477-630-11  
: Sequence 11, Application US/08477630  
: Patent No. 5721128  
: GENERAL INFORMATION:  
: APPLICANT: DEMEER, PHILIPPE  
: APPLICANT: AMORY, ANTOINE  
: TITLE OF INVENTION: PULULIANASE, MICROORGANISMS WHICH  
: TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
: TITLE OF INVENTION: PULULIANASE AND THE USES THEREOF

NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
City: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
Zip: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,630  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-42  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-1850  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 928 amino acids  
TYPE: amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-630-11

Query Match 96.6%; Score 4888.5; DB 1; Length 928;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 926; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 30 DGNNTTIIIVHFCPCAGDYQWPWSLMMMPKDGGAETDFNOPADSGAASADIPGNPSQVG 89  
Db 1 DGNNTTIIIVHFCPCAGDYQWPWSLMMMPKDGGAETDFNOPADSGAASADIPGNPSQVG 60  
QY 90 IIVRTQDMTKVDSADRYIDLSKNGEVLVEGNSQIFYNKEKAEDAAAPASNAVLADSNQ 149  
Db 61 IIVRTQDMTKVDSADRYIDLSKNGEVLVEGNSQIFYNKEKAEDAAAPASNAVLADSNQ 120  
QY 150 VLVLKSPPLTIGEGXSGFTVHDDTANKNDIPYTSVKDASLGDVYAVLAGTQHIFFGSDW 209  
Db 121 VLVLKSPPLTIGEGXSGFTVHDDTANKNDIPYTSVKDASLGDVYAVLAGTQHIFFGSDW 180  
QY 210 APDNHSTLLKAVTNKLVQFSGDLPEGNQYQVVALNDSNN-SYPSDNIINTVPAGAHVT 268  
Db 181 APDNHSTLLKAVTNKLVQFSGDLPEGNQYQVVALNDSNN-SYPSDNIINTVPAGAHVT 240  
QY 269 FSYIPSTHAYVDTINNPADLOVESGKVTDLVTVTLGEDPVSHTLSIQTDGYAKQVITP 328  
Db 241 FSYIPSTHAYVDTINNPADLOVESGKVTDLVTVTLGEDPVSHTLSIQTDGYAKQVITP 300  
QY 329 RNVLNNSQYYSGGDLGNTYQKATTFKVMAPSTOVAVLLYDSATGSVTKIVPMTASGH 388  
Db 301 RNVLNNSQYYSGGDLGNTYQKATTFKVMAPSTOVAVLLYDSATGSVTKIVPMTASGH 360  
QY 389 GWEATVNONLENNYVYMEVYGCGSTRPAVDPAATAPNCTRCMIYDLAKTDPAGWNSD 448  
Db 361 GWEATVNONLENNYVYMEVYGCGSTRPAVDPAATAPNCTRCMIYDLAKTDPAGWNSD 420  
QY 449 KHITPKNIEDEVYIEMVDROFSIDPNSGMMKKGYLALTEKGTGPDVVKGTGIDSLQOLG 508  
Db 421 KHITPKNIEDEVYIEMVDROFSIDPNSGMMKKGYLALTEKGTGPDVVKGTGIDSLQOLG 480  
QY 509 ITHVOLMPVFASNSVDETPQDMNGGIDPRNYDVPBGQYATNANGNARIKFEKEMVLSLH 568  
Db 481 ITHVOLMPVFASNSVDETPQDMNGGIDPRNYDVPBGQYATNANGNARIKFEKEMVLSLH 540



```

QY      569 REHGVAMDVYVNHHTATOTSDSDKIVPEXYRTMIOVITPTDOYLEMKIAERPMVOKF   628
        |||||||
Db       541 REHGVAMDVYVNHHTAQTQSDDKIVPEYTRTMQVITPTDOYLEMKIAERPMVOKE   600
QY      629 IIDSLLKAYWNEHYHIDGFRFDLMALGKDTMSKAASLEHAINPGIALYGEPMVTGTSALPD  688
        |||||||
Db       601 IIDSLKAYWNEHYHIDGRFPLMALLGKDVTMSKAASLEHAINPGIALYGEPMVTGTSALPD  660
QY      689 DOLLTKRAGKGAVAFNDLRLAALLGNVPDSSAGCFATGATGLTDAIKNGVEGINFT    748
        |||||||
Db       661 DOLLTKRAGKGAAVFANDRLAALLGNVPDSSAGCFATGATGLTDAIKNGVEGINFT    720
QY      749 SSPEGTYNTYSHDNVTILMKI.LSNPNSEADRIKMDELAOAVMTSGCVFMGGEGEM   808
        |||||||
Db       721 SSBETINYTSHDNTLTMDKI.LSNPNSEADRIMDELAOAVMTSGCVFMGGEGEM   780
QY      809 LRXKGDNSNYMGDAVNEFDSMRKAOYPDVFNYTSGLIHLRLDPAPRMTTANEINSHL   868
        |||||||
Db       841 QFLNSPENTVAVELTDNVNNDKGNITIVYNPKRYATATINLP SGKVALNATSGKRGESTL   900
QY      929 GQAGSVQVPGISMMILHQEVSPPDHCK   956
        |||||||
Db       901 GQAGSVQVPGISMMILHQEVSPPDHCK   928

RESULT      3
US-08-472-293-11
; Sequence 11, Application US/08472293
; Patent No. 5731174
; GENERAL INFORMATION:
; APPLICANT: DEWEER, PHILIPPE
; APPLICANT: AMORY, ANTOINE
; TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH
; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS
; TITLE OF INVENTION: PULULANASE AND THE USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K Street, N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,293
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilhelm F. Gadlano, Esq.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-44
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEO ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-293-11

```

Query Match	96.6%	Score 4888.5	DB 1	Length 928
Best Local Similarity	99.8%	Fred. No. 0		
Matches 926	Conservative 1	Mismatches 0	Indels 1	Gaps 1
QY 30	DGNNTTIIIVHYCCPAGDYPQPMSLMMMPKDGCAEYDFNQPADSPFGAVASADIPGNPSQVC	89		
DB 1	DSNTTTIIVHYFCPAGDYPQPMSLMMMPKDGCAEYDFNQPADSPFGAVASADIPGNPSQVC	60		
QY 90	IIVPTQDMTKDVSADRYIDLSKGNFVWLVEGNSQIFTEKKADEAKKAPVSNAYLZASNQ	149		
DB 61	IIVPTQDMTKDVSADRYIDLSKGNFVWLVEGNSQIFTEKKADEAKKAPVSNAYLZASNQ	120		
QY 150	VLYKLSQPLTJGEGKSGFTVHDDTANKDIPVTSYVDASLSLQGVYAVLACTPGHITGGSGW	209		
DB 121	VLYKLSQPLTJGEGKSGFTVHDDTANKDIPVTSYVDASLSLQGVYAVLACTPGHITGGSGW	180		
QY 210	APDNHSTLLKKTATNNLYOFSGDLPGCANQYKALANDSMNN-SYPSDNIULTYPACGAHYT	268		
DB 181	APDNHSTLLKKTATNNLYOFSGDLPGCANQYKALANDSMNNSPSPDNIULTYPACGAHYT	240		
QY 269	PSYIDPSTHAVIDTINNPNADLOVESGVKTDLVTVTLGEDPDVSHTLSTIOTDGYQAKQVIP	328		
DB 241	PSYIDPSTHAVIDTINNPNADLOVESGVKTDLVTVTLGEDPDVSHTLSTIOTDGYQAKQVIP	300		
QY 329	RNVLLSSQYYSGGDLGNNTYQKATPEKMAVSTSOVNVLLVDSATGVSATKIVPMTASGH	388		
DB 301	RNVLLSSQYYSGGDLGNNTYQKATPEKMAVSTSOVNVLLVDSATGVSATKIVPMTASGH	360		
QY 389	GWEATVNTQNLNNMYMYEVTGQSGSTRPAVDYATATAIPNGTRGMIIVDLAKTPDAGMNSD	448		
DB 361	GWEATVNTQNLNNMYMYEVTGQSGSTRPAVDYATATAIPNGTRGMIIVDLAKTPDAGMNSD	420		
QY 449	KHITTKNIIEDEYIYEMDVRDESIDPNSGKNKNGKTLIALTEKTKRGPDNKTKGIDSLKQLG	508		
DB 421	KHITTKNIIEDEYIYEMDVRDESIDPNSGKNKNGKTLIALTEKTKRGPDNKTKGIDSLKQLG	480		
QY 509	ITHVQLMPVFASSNVDEPTQDNNGGYDPRANDVEGQATNANGNARIKEFEKWLSLH	568		
DB 481	ITHVQLMPVFASSNVDEPTQDNNGGYDPRANDVEGQATNANGNARIKEFEKWLSLH	540		
QY 569	REHIGVNMDDVYVNHPTAOTISDPDKIVPEYIYRTMIQVITPTDOVLEMLKLAERPVAQKE	628		
DB 541	REHIGVNMDDVYVNHPTAOTISDPDKIVPEYIYRTMIQVITPTDOVLEMLKLAERPVAQKE	600		
QY 629	IIDSILKMYWNEVHIDGFERPDLMAILGKDTMSRAASELHAINCGIALYGEPMVGGTSALPD	688		
DB 601	IIDSILKMYWNEVHIDGFERPDLMAILGKDTMSRAASELHAINCGIALYGEPMVGGTSALPD	660		
QY 689	DOLLTRGAKGKGVAVFNQNLNRNALDGNINFDSSAOCFAFAGTGLDIAINGEGSINDPT	748		
DB 661	DOLLTRGAKGKGVAVFNQNLNRNALDGNINFDSSAOCFAFAGTGLDIAINGEGSINDPT	720		
QY 749	SSPGETINVTVSHDNYTLMDKIALSNPNDSSEADRIKMDLEAOAVVMTSGCVPFMOCGEEM	808		
DB 721	SSPGETINVTVSHDNYTLMDKIALSNPNDSSEADRIKMDLEAOAVVMTSGCVPFMOCGEEM	780		
QY 809	LKKXGNDNSYAGDAVNEFDMSSRAQYQDVYVNSYSGTLHLIDHPAFMTTANETNSHL	868		
DB 781	LKKXGNDNSYAGDAVNEFDMSSRAQYQDVYVNSYSGTLHLIDHPAFMTTANETNSHL	840		
QY 869	QPLNSPENTVAYELTDHVNRDKMGNTIIVVYNNPKTIVATINLPSGKMAINATSGKVGEESTL	928		
DB 841	QPLNSPENTVAYELTDHVNRDKMGNTIIVVYNNPKTIVATINLPSGKMAINATSGKVGEESTL	900		
QY 929	GQAEQSVQYVPGISMMILHQEVSPDHGKK 956			
DB 901	GQAEQSVQYVPGISMMILHQEVSPDHGKK 928			

RESULT 4  
US-08-474-545-11  
: Sequence 11, Application US/08474545

Patent No. 5736375  
GENERAL INFORMATION:  
APPLICANT: DEWEER, PHILIPPE  
APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOPER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,545  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 928 amino acids  
TYPE: amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-474-545-11

Query Match 96.6% Score 4888.5 DB 1: Length 928;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 926: Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
1;  
30 DCGNTTIIIVHFCPCADYQPMWLMMPKDGCAEYDFNOPADSGAASADIPGNPQVG 89  
1 DGNNTTIIIVHFCPCADYQPMWLMMPKDGCAEYDFNOPADSGAASADIPGNPQVG 60  
90 IIVTQMTKDVADRYIDLSKGNVWLVESNSQIFNEKDAEDAAKPAVSNAIPLDASNQ 149  
61 IIVTQMTKDVADRYIDLSKGNVWLVESNSQIFNEKDAEDAAKPAVSNAIPLDASNQ 120  
150 VLVKLSQPLTLEGSGSGFTVHDDTANKDIPYTSKDSLGDVTAVALAGTQHFEGSDW 209  
121 VLVKLSQPLTLEGSGSGFTVHDDTANKDIPYTSKDSLGDVTAVALAGTQHFEGSDW 180  
210 APDNHSTLLKRVNTNLYQFSGDLPEGNVQYKVALNDSMN-SYPSDNIINLTPAGAHVT 268  
181 APDNHSTLLKRVNTNLYQFSGDLPEGNVQYKVALNDSMN-SYPSDNIINLTPAGAHVT 240  
269 FSYIPSTHAYVDTINNRAADQVSSGKTDLYVTLEDDPVSHTLSIQDGYAKQVYIP 328  
241 FSYIPSTHAYVDTINNRAADQVSSGKTDLYVTLEDDPVSHTLSIQDGYAKQVYIP 300  
329 RNVLNSQYYSGGDLGNTYQKATTFKVMAPSTQYVNLVLYDSATGSVTKIVPMTASGH 388  
301 RNVLNSQYYSGGDLGNTYQKATTFKVMAPSTQYVNLVLYDSATGSVTKIVPMTASGH 360  
389 GWEATVNONLENNYMYEVTGSGSTRADVAPATAIAPNCRGMIVDLAKTDPAGNSD 448  
361 GWEATVNONLENNYMYEVTGSGSTRADVAPATAIAPNCRGMIVDLAKTDPAGNSD 420

449 KHITPKNIEDEVIEMKVDYDFSDPNSGMKNGKYLALTEKGTGPDNVATGIDSLKQLG 508  
421 KHITPKNIEDEVIEMKVDYDFSDPNSGMKNGKYLALTEKGTGPDNVATGIDSLKQLG 480  
509 ITHVQMLPVASNSVDLTPDQDMWGYDPNRYDVPEQGYATNAGNARIIEFKEMVSLH 568  
481 ITHVQMLPVASNSVDLTPDQDMWGYDPNRYDVPEQGYATNAGNARIIEFKEMVSLH 540  
569 REHIGVMDVYVNTFPAQISDFDKIPEYRYRTMIOVITPDQVLEMKLAERPYNOKF 628  
541 REHIGVMDVYVNTFPAQISDFDKIPEYRYRTMIOVITPDQVLEMKLAERPYNOKF 600  
629 IISLAKTVANEYHIDGRFDLMLLKGDTMSKASELHAINPGIALYGEPTGTSLAPD 688  
601 IISLAKTVANEYHIDGRFDLMLLKGDTMSKASELHAINPGIALYGEPTGTSLAPD 660  
689 DQLTGKAGKMGVAVPNDLNRNALDGNVPSQAGFATGATGLDIAKNGVEGSIINDFT 748  
661 DQLTGKAGKMGVAVPNDLNRNALDGNVPSQAGFATGATGLDIAKNGVEGSIINDFT 720  
749 SSPGETINVTSHDNTYLMKIALSNPNDSADRIKMDLAQAVVMTSGVPMQGGSEEM 808  
721 SSPGETINVTSHDNTYLMKIALSNPNDSADRIKMDLAQAVVMTSGVPMQGGSEEM 780  
809 LRKXGMDNSTNAGDAVNEPMSRKAQYPDVFNTYSGLIHLRDPAPRTTANETINSHL 868  
781 LRKXGMDNSTNAGDAVNEPMSRKAQYPDVFNTYSGLIHLRDPAPRTTANETINSHL 840  
869 QFLSPENTVAYELTDVHNKDKMGNIIIVYNPKTVATINLPSGKMAINATSGVGRSTL 928  
841 QFLSPENTVAYELTDVHNKDKMGNIIIVYNPKTVATINLPSGKMAINATSGVGRSTL 900  
929 QGAEQSVQVPGISMWILHQEYSPDHGRK 956  
901 QGAEQSVQVPGISMWILHQEYSPDHGRK 928

RESULT 5  
US-08-478-341-11  
Sequence 11, Application US/08478341  
Patent No. 5817498  
GENERAL INFORMATION:  
APPLICANT: DEWEER, PHILIPPE  
APPLICANT: AMORY, ANTOINE  
TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOPER GILSON & LIONE, P.C.  
STREET: 2000 K Street, N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,341  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-45  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 928 amino acids  
 TYPE: amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-478-341-11

Query Match 96.6%; Score 4888.5; DB 2; Length 928;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 926; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

30 DGNNTTIIIVHFCPPADYQPMSTLMMPKDGGAGYEYFNPADSFQAVASADIFGNSQVQ 89  
 1 DGNNTTIIIVHFCPPADYQPMSTLMMPKDGGAGYEYFNPADSFQAVASADIFGNSQVQ 60  
 90 IIVRTQDMTKDVSADRYIDLKSGNEVWLVEGNSQIFYNKEDADDAKPAVSNAYLDAASQ 149  
 61 IIVRTQDMTKDVSADRYIDLKSGNEVWLVEGNSQIFYNKEDADDAKPAVSNAYLDAASQ 120  
 150 VLVKLSQPLTLEGSGSGETVHDDTANKDIPVTSVKDASLGQDVTAVLAGTFQHIFGSDM 209  
 121 VLVKLSQPLTLEGSGSGETVHDDTANKDIPVTSVKDASLGQDVTAVLAGTFQHIFGSDM 180  
 210 APDNSTLTKKVTNNLYQFSGDLPEGNTQYKVALDMSNN-SYSPDNINLTVPAGAHYT 268  
 181 APDNSTLTKKVTNNLYQFSGDLPEGNTQYKVALDMSNNPSYSPDNINLTVPAGAHYT 240  
 269 FSYISTHAYVDTINNPADLQVESGKTDVTLTGEDPDVSHLTSTQTDGQAKQVTP 328  
 241 FSYISTHAYVDTINNPADLQVESGKTDVTLTGEDPDVSHLTSTQTDGQAKQVTP 300  
 329 RNVLNLSQYYSGGDLGNTYTKATTFKVMAPTSTOVNVLTYDSATGVSYTKIVPMASGH 388  
 301 RNVLNLSQYYSGGDLGNTYTKATTFKVMAPTSTOVNVLTYDSATGVSYTKIVPMASGH 360  
 389 GWEMATVNMNLNMTYETVGGSTRTAVDPYATAIAPNGTRGMIVLAKTDPAWMSD 448  
 361 GWEMATVNMNLNMTYETVGGSTRTAVDPYATAIAPNGTRGMIVLAKTDPAWMSD 420  
 449 KIITPKNIEDEYIEMDVDFSIDPNSGKMKKYLALTEKGTGGPDNKTGIDSLKQIG 508  
 421 KIITPKNIEDEYIEMDVDFSIDPNSGKMKKYLALTEKGTGGPDNKTGIDSLKQIG 480  
 509 ITHVQLMPFVAFASNYDETDPTQDNNGYPRNDVPEGQYATNANGNARIKEKEWYLSH 568  
 481 ITHVQLMPFVAFASNYDETDPTQDNNGYPRNDVPEGQYATNANGNARIKEKEWYLSH 540  
 569 RHHIGVNMNVYNNHTFAQISDPDKIVPEYRTYRMIVQIIPTDQVLEMKLAERPVQKF 628  
 541 RHHIGVNMNVYNNHTFAQISDPDKIVPEYRTYRMIVQIIPTDQVLEMKLAERPVQKF 600  
 629 IIDSLLKYWNEVHIDGFRDLAALLGKDTMSKASSELHINFGIALYGEPMGTGTSALPD 688  
 601 IIDSLLKYWNEVHIDGFRDLAALLGKDTMSKASSELHINFGIALYGEPMGTGTSALPD 660  
 689 DOLLTKGAQKMGVAVFNDNLNADLGNVFDSSAOGFATGATGLDAIKNGVGSINDFT 748  
 661 DOLLTKGAQKMGVAVFNDNLNADLGNVFDSSAOGFATGATGLDAIKNGVGSINDFT 720  
 749 SSPGETINTVTHDNYTLMDKIALSNPNDSEADRIKMDLAQAVVMTSGCVFPMOGGEM 808  
 721 SSPGETINTVTHDNYTLMDKIALSNPNDSEADRIKMDLAQAVVMTSGCVFPMOGGEM 780  
 809 LKXKGNDSYVAGDAVNEFDMRSKRAQYPDVFNYSGLIHLRLDHPAFMTTANETNSHL 868  
 781 LKXKGNDSYVAGDAVNEFDMRSKRAQYPDVFNYSGLIHLRLDHPAFMTTANETNSHL 840  
 869 QFLNPEPNTVAVELTDVHKDKMGNIIVYNNPKTATINTLPSGKMAINATSGKVESTL 928  
 841 QFLNPEPNTVAVELTDVHKDKMGNIIVYNNPKTATINTLPSGKMAINATSGKVESTL 900  
 929 GOAEGSVQVPGISMMILHQVSPDHCK 956

DB 901 GOAEGSVQVPGISMMILHQVSPDHCK 928

RESULT 6

US-08-996-733-11

Sequence 11, Application US/08996733  
 Patent No. 6074854  
 GENERAL INFORMATION:  
 APPLICANT: DEMEER, PHILIPPE  
 APPLICANT: AMORY, ANTOINE  
 TITLE OF INVENTION: PULLULANASE, MICROORGANISMS WHICH  
 TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International, Inc.  
 STREET: 925 Page Mill Road  
 CITY: Palo Alto,  
 STATE: CA  
 COUNTRY: U.S.A.  
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,733  
 FILING DATE: 23-DEC-1997

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/472,293  
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/174,893  
 FILING DATE: 28-DEC-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: BE 09301278  
 FILING DATE: 19-NOV-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: BE 09300744  
 FILING DATE: 15-JUL-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: BE 09201156  
 FILING DATE: 28-DEC-1992

ATTORNEY/AGENT INFORMATION:  
 NAME: Debra J. Glazier, Esq.  
 REGISTRATION NUMBER: 33,888  
 REFERENCE/DOCKET NUMBER: GC446C1-US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 846-7620  
 TELEFAX: (650) 845-6504

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 928 amino acids  
 TYPE: amino acids  
 TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-996-733-11

Query Match 96.6%; Score 4888.5; DB 3; Length 928;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 926; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

30 DGNNTTIIIVHFCPPADYQPMSTLMMPKDGGAGYEYFNPADSFQAVASADIFGNSQVQ 89  
 1 DGNNTTIIIVHFCPPADYQPMSTLMMPKDGGAGYEYFNPADSFQAVASADIFGNSQVQ 60  
 90 IIVRTQDMTKDVSADRYIDLKSGNEVWLVEGNSQIFYNKEDADDAKPAVSNAYLDAASQ 149  
 61 IIVRTQDMTKDVSADRYIDLKSGNEVWLVEGNSQIFYNKEDADDAKPAVSNAYLDAASQ 120

QY	150	VLKLSQPLTLGGXSGFVYHDTAKRDLPRMVSVDASLGDQVTPAVLACTGPHLFGSGDM	209
Db	121	VLKLSQPLTLGGXSGFVYHDTAKRDLPRVSVDASLGDQVTPAVLACTGPHLFGSGDM	180
QY	210	APDNHSTLLKKTNNLKYRSGDLPEGNYQYKVALNDSMNN-SYPSDNLINLTPAGAHVT	268
Db	181	APDNHSTLLKKTNNLKYRSGDLPEGNYQYKVALNDSMNNPSYPSDNLINLTPAGAHVT	240
QY	269	FSYTPSHAVYDTINPNNDLQYEGSVKTRDLVTYTLGDEPDVSHLSTOTGQYAKQYIP	328
Db	241	FSYTPSHAVYDTINPNNDLQYEGSVKTRDLVTYTLGDEPDVSHLSTOTGQYAKQYIP	300
QY	329	RNVLNSQYIYSGDDLQNTYTKATTFKVMAPTSTQVNVLLYDSATGSVTKIYVPTASGH	388
Db	301	RNVLNSQYIYSGDDLQNTYTKATTFKVMAPTSTQVNVLLYDSATGSVTKIYVPTASGH	360
QY	389	GWEATYVNOULEMYIYMYEYTGGSRTFVADYATAPALAPNGTRGATVLDATDPAGNSD	448
Db	361	GWEATYVNOULEMYIYMYEYTGGSRTFVADYATAPALAPNGTRGATVLDATDPAGNSD	420
QY	449	KHLPKRIEDEVYENDVDRFESIDPMSGKNNKGKYLALTEKTKRDPDNNYKGISLRKG	508
Db	421	KHLPKRIEDEVYENDVDRFESIDPMSGKNNKGKYLALTEKTKRDPDNNYKGISLRKG	480
QY	509	ITVYOLMPVFNASVDETFDQDNMGYDPRNVDYVEGQYATNANGNARIKEFEKXVLSLH	568
Db	481	ITVYOLMPVFNASVDETFDQDNMGYDPRNVDYVEGQYATNANGNARIKEFEKXVLSLH	540
QY	569	REHGVMMDDVYVNHTRATQISDFKIVPEYRTMIOVYIIPTDQYLEMKLAEKRPVQKF	628
Db	541	REHGVMMDDVYVNHTRATQISDFKIVPEYRTMIOVYIIPTDQYLEMKLAEKRPVQKF	600
QY	629	IIDSLLKAYNEHYHIDGFRDLALLGKDTMSKAASELHAINGIALYGEPTWGTSLALPD	688
Db	601	IIDSLLKAYNEHYHIDGFRDLALLGKDTMSKAASELHAINGIALYGEPTWGTSLALPD	660
QY	689	DOLLTKRQKGMCAVAFPNNDLRYALDGNFPPDSAGCFATGALGLDAIKNGYEGSINFT	748
Db	661	DOLLTKRQKGMCAVAFPNNDLRYALDGNFPPDSAGCFATGALGLDAIKNGYEGSINFT	720
QY	749	SSPGETINVTSHDNTYLMKIALISPNPNDSEADRIKMDLQAAVMTSGVYFMQSGEM	808
Db	721	SSPGETINVTSHDNTYLMKIALISPNPNDSEADRIKMDLQAAVMTSGVYFMQSGEM	780
QY	809	LRRXGGNDNSYNAGDAVNEFDSRKQYQDYVFNYSGLLHLHLDPARMTTANINSHL	868
Db	781	LRRXGGNDNSYNAGDAVNEFDSRKQYQDYVFNYSGLLHLHLDPARMTTANINSHL	840
QY	869	QFLNSPNTAAVELTDVNVNKGKNIIVVYNPKNVTATINLPSGKMAINATSGKXGSETL	928
Db	841	QFLNSPNTAAVELTDVNVNKGKNIIVVYNPKNVTATINLPSGKMAINATSGKXGSETL	900
QY	929	GOAEGSVQVPGISMMILHOFVSPDHGK 956	
Db	901	GOAEGSVQVPGISMMILHOFVSPDHGK 928	
RESULT 7			
US-08-860-339-18			
Sequence 18, Application US/08860339			
Patent No. 6117665			
GENERAL INFORMATION:			
APPLICANT: Kossmann, Jens			
APPLICANT: Eimermann, Michael			
APPLICANT: Virgin, Ivar			
APPLICANT: Renz, Andreas			
TITLE OF INVENTION: DNA MOLECULES CODING FOR DEBRANNING ENZYMES DERIVED			
TITLE OF INVENTION: FROM PLANTS			
FILE REFERENCE: AGREVO-6			
CURRENT APPLICATION NUMBER: US/08/860, 339			
CURRENT FILING DATE: 1997-11-25			
EARLIER APPLICATION NUMBER: DE P4447387.7			
EARLIER FILING DATE: 1994-12-22			

```

:      NUMBER OF SEQ ID NOS: 24
:      SOFTWARE: PatentIn Ver. 2.0
:      SEQ ID NO 18
:      LENGTH: 964
:      TYPE: PRF
:      ORGANISM: Spinacia oleracea
:      US-08-860-339-18

```

Query Match	15.2%;	Score 771;	DB 4;	Length 964;
Best Local Similarity	26.0%;	Pred. NO. 8e-52;		
Matches 260;	Conservative 133;	Mismatches 320;	Indels 286;	Gaps 39

QY	196	LACFOHIFCGSDAPDN-----HSTLKKVTN-----	2233
Dp	10	LASSFHHHPNLRPLPNNFNFTKLPVNSPFAIGSSSRSHSPKDKDSCFCCSNAVEY	69
QY	224	-----NLVQFSGDDLPECNQY-----KVALANDSMNYS	252
Dp	70	GSASSVQSCELOGLSNCRAVWPFSKYFANNVNDIGNSYLLFAFSKTALEKFTDAGIEGY	128
QY	253	SDNLNLVPAGG--AHT--FSYIPRSHAVYDPIINNNAOLOVESGKTDLYVTIGEDP	308
Dp	129	DVKIKLKDQGLPRANYTEKPHIRG---YSAFKAP-ATLVDSDSLCOLAAVAASADG	183
QY	309	DVHSTLSIOTDGYOAKOVIPRNVJNNSOYVYSGDDLNTTOKATTFKVAAPSTOVNVL	368
Dp	184	ACRNATGLOLRG-----VIDELYSYDG-PICGAVESENTISLYMAPTAOAVSAS	231
QY	369	LYDSATGSYKTIYPMTASGHCWPEATYONOLENNYTYEVT--GQSGTRP---ANDPYA	422
Dp	232	IFKPSGSEPLQTVOLIESNGWASVAPRMEGCGYVYETIYVHSHSTRLEKSFALDPA	291
QY	423	TAIPAGTRGMIVDLAK--TDPAAMNSDKHTPKNIDE-----YIYENDVDFG	470
Dp	292	RGISADVKRTLLADLSEBTLKPEGM-----ENLADKERPHILSPDISLSYELHTRDFS	343
QY	471	-----IDPNSGMKKKGYTLALTEKGTGAPDNVKTGIDSLKQLGITHVQMLPVPASVSD-	524
Dp	344	AYDULVHPDL-----RGCVLAFTSDQSAGVNH--LEKLSAAGITGHLLPSPFOAEVDD	395
QY	525	-----ETDLPQD-----NMGYDPRIYVDEPCOYATNANGNA	555
Dp	396	DKKKMKFVDKRFETLRPDSEEOAQITLARDECGYMWGYNPVLMGPKPSYATDPNGPC	455
QY	556	RIKEFEKEVJLSLAREHIGKMDVVYNNHTFATQISD---FDKIVPEYVYFTMIQVILPTD	611
Dp	456	RIIEFRKKVQALNRIGLRVYLDVYVNNHLNSSGSPDDNSYVDKIPYGYLLRR-----DND	509
QY	612	QVLEBMK-----LXKERPMYOKFIIDSLKYYWNEVHIDGCFPDLMALLGKXTMSKASELH	666
Dp	510	GALNSTCVDNTASEHNEVERLLIIDLKHMAVYNNKVGCFRERDMLGHITMKHTMYKATMQLQ	569
QY	667	A1N-----PGIALYGEPTTGTGTSALPDDQLLTICKAQ--KOMGYAVPNDLRNA-LD	714
Dp	570	GLSKNIDQVEGSSYIYXGEGMDPEVA--NNARGVNMSQLNGTGTGCFPNDRIDRAVLG	627
QY	715	GNVVDSS-AQGFATGAR-----GLTAKNGVGSIND-----	746
Dp	628	GGPGRPLQDGYVGLSLQPDNDHDSGRKANADRLAAVAKHIDQVGMAGNLRDVIYLTNCDCG	687
QY	747	-----FTSGPGETIYVYTHSHDVTYLMDKILASNPNDSEAD-RIKKDEL	788
Dp	688	KQVNGSEYTYTGGTRPVGYAMOPITITIIYVSAHDETELFDIYSLKTPIYTYVDERCKRVNHL	747
QY	789	AAQAVVMTSGVPEFMQGEEMLRKKGNDNSYNAGDAVNEFMS-----	831
Dp	748	ATSLIALSQGIPFPHAGDELLRSKSLDRDSYNSGDMFNRLDEFYSNMMGCVGLPRKDNE	807
QY	832	-----RKAOYD-----VFYVYSGLLHLRLDHPAFRPTTANELNSHLQFANSP	874
Dp	808	SNMPLIKRILANPSYKDKDKNHI1AAVENFNLLQIRSSSPLEFR:SKSDIEDVRFAPNNY	867

OY 875 EN---TVAYELTD-HVN-----KDKMGNTIIVYVNPKNKVTATINDPSGK-----NAI 916  
Db 868 PSMIGLIMSIEDGHAGAPLSQIDPKFOYIVVLIINQPIETKFNPNDLRAKSLQHPV 927  
OY 917 NATSGK--VGEISTLGAEGSVQVPGISMMILHGEVSPDH 953  
Db 928 QSTGDIYVAKESKYEPSGCTTIPKSTAVP---VEPRH 963

## RESULT 8

US-08-961-083-72

Sequence 72, Application US/08961083

Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.083

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 72:

SEQUENCE CHARACTERISTICS:

LENGTH: 618 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-961-083-72

Query Match 10.7%; Score 543; DB 4; Length 618;

Best Local Similarity 30.3%; Pred. No. 2,9e-34;

Matches 149; Conservative 82; Mismatches 169; Indels 92; Gaps 18;

OY 507 LGITHVQLMPVYASNSVE-----TDPTODNMGYDRNVDVPGOYATN-ANGN 554  
Db 1 LGVTHIQLLPVLSYFYFNELKHKHERLSDYASSNSNVMQCYDONFSLTGMSSDPKNE 60  
OY 555 ARIKEFEKEMVSLAREHIGVNDVVYVNTFATQISDFDKIPEYVYRTMIOVYIIPDQVL 614  
Db 61 KRIAEFKNLINIEHRRKGALIDVYVNH7--AKVDLFEDLEPNYH-FMDADGTPRTSG 117  
OY 615 EMKLAARPMYOKFTIDSLKYVNVNEYHIDGFRFDLMALLGRTKMSKASELAINGIAL 674  
Db 118 GRLGTTHHMTKRLILDISIKYLVDTYKVGFRDMGHDAAISIEAYKARALNPMLIM 177  
OY 675 YGEPTGTSTALPDQLLTGKAOKG-----GVAVFNDLRLNALDGNVDFSSAOGFATGA 729  
Db 178 LGECH---RTVAGDENMPTKADQDMKHTDVAVFSDDIRNNLNSGYNNEGQPAFITG 234

OY 730 TGLTDAI-KNVEGSIINDFTSSPGETINYVTSNDVTLMDKIALS---NPNDSE--ADRI 783  
Db 235 KRDVNTIEKNIQAQTNFEADSPGDVIGYIAHNDLLEFDIAQSIKDPKSAENVAEIH 294  
OY 784 KNDLGAQVAVMTSGVPPFNGGEE-----MRXKGN-- 815  
Db 295 RRLRGNLMLVLAQSTPPLHSQGEYGRFKOPRDPAPKTPVADKVPNNKSHLLRDXGNPF 354  
OY 816 -----DNSYNAGDAVNEFEGRKAO---YPDVF---NYSGLIHLRDPHAFRRMTANE 863  
Db 355 DYPIFIHDSYSSDAVKNKDWTKATDKGKAYPENKSRDYMKGILALRGSTDAFRLKSLD 414  
OY 864 INSHLOFLNSP-ENTVAYELTDHVNKDKWNTIIVYN---PKVTATINLPSGKAINATF 919  
Db 415 IKDRVHLITVPGQNGYERE-----DVVIGQIYAPNGDIYAV-----FVNND 456  
OY 920 SGKVGESTLGA 931  
Db 457 E-KAREFNLTGA 467

## RESULT 9

US-08-476-519-11

Sequence 11, Application US/08476519

Patent No. 5750876

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Krohn, Bradley M.

TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F

STREET: 700 Chesterfield Parkway No. 5750876th

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476.519

FILING DATE:

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/281902

FILING DATE: 28-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Bonner, Grace L.

REGISTRATION NUMBER: 32,963

REFERENCE/DOCKET NUMBER: 38-21(13577)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-7286

TELEFAX: (314)537-6047

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 746 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-476-519-11

Query Match 10.1%; Score 509; DB 1; Length 746;

Best Local Similarity 25.5%; Pred. No. 1.8e-31;

Matches 191; Conservative 91; Mismatches 290; Indels 176; Gaps 27;

OY 344 LGNTY--TOKATTFKYMATPSTQVNVLLYDSATG--VTKRIPTMASGHWAEATV-NON 398  
Db 116 LGNTY--TOKATTFKYMATPSTQVNVLLYDSATG--VTKRIPTMASGHWAEATV-NON 398

```

Db      8 LGARDAQAQNLAFRVYSRSATREVEFLYKKNPTSGSEVARLALSKDPATQVMSLSLPTST 67
Oy      399 LENM-----TYMEVYTG-----GCS-----TRTAVDPATA 424
Db      68 1KNVYGITCAVAYYGRRAWGPMWPDAAWTKGSATGFSVDVDMACNRPKLLDIPYARE 1277
Oy      425 IAPNCTGMIYD--LAKTDPAGMNSDKHI-----TPKNIEDEVYEM 464
Db      128 ISQDPNTATCADGTTIATGCAHRRKDSGLCSKSIATAADATVSGSKPTFALDEVIYEV 187
Oy      465 DYRDESIDPNSGMKNKKGYTLALTEKGT-KGPDNVKKTGISLKOIGITHVOLMPVFAS--- 520
Db      188 HVR-----GLTRNDSDVPAERGGTYKGAARKAA---ALAAAGTAYAVEFLPYOEQRND 236
Oy      521 -NSVDETPDPTQDN-MGYVPRAYDYDEGQYATNANGNARIKEFEKEMVLSLREHIGVMYD 578
Db      237 ONDVDPNSTACDNTWGMVTLTTFAPDRKRYATDKSAGGPTREKMAVAFHDAGIKYITDY 296
Oy      579 VYNIIT-----FATQISDFKATVPEYYRTIMQVITPTDY-LEMKLXAERPVM 625
Db      297 YVNHTEGEGCPMSGTGLDSVYMLLSFRGLDNPAYYSLSSDYKYPMDNTGVCGNVTRHP1A 356
Oy      626 QKFIIDSLKYVNEHIDGFRPDLMALIGK-----DTMSKAASEL---H 666
Db      357 QNLVYDSLAYVRDALGVDFRFDLASVYLGNSQHGFCTNFDKNDSGNALNRIIVELPPRA 416
Oy      667 AINPGIALYGERPWT-GCTSALRPDDOLLTKGAOKG---MGVAVFENDNIRNMLDGNVFPSSA 722
Db      417 AGCAGACADTIAEPWALIGNS-----YVGGFPAGNAEMNGLYRDALRKKNKLG 465
Oy      723 QGFATGATGLDAIKNGVEGSIINFTSSPGTITVYVSHDNTITMDKIALSNRPDSE--- 779
Db      466 ETVTPTGLATRPAGSNDLYG---DDGRKPWHSINFVAHAGFPTLNDLYAVNDKONNDPWP 522
Oy      780 -----ADRIKMDLAQAVVMTSGVYPMQCGCEMLRXKGCNDNSY 819
Db      523 YGPSGCGGDHMLSNMGGSIYVAEQKRAARTGALLMLLSAGVPMITGGEDALRTQGNNTY 582
Oy      820 NAGDAVNEFDNSRKAQYRDPVYNYSGLIHLRLDHPAFR---MTTANEINSHLO----- 869
Db      583 NLDASAMNLYMSRSRLLEADHETVYKRLIAFRAKAPALRPANFYSASDNTGNVMEOLRWFK 642
Oy      870 -----PLNSPEN-IVAAELDIHVKKDKGNIIVYVNTKRVATINLP---SGKMAI 916
Db      643 PDGAQADATATRCADNHALAKRIQSEFCDSASATYYAIVNGMGAUVFKLPWETGKQWY 702
Oy      917 NATSGKVESTLCOAEGSVQVPGISMMI 944
Db      703 RYTD---TATWNEGPNAVVALPGSETIL 726

RESULT 10
PCT-US95-09323-11
Sequence 11, Application PC/TUS9509323
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
NUMBER OF SEQUENCES: 11
TITLE OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09323
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/281902
FILING DATE: 28-Jul-1994
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 746 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09323-11

```

Query Match	10.1%;	Score 509;	DB 5;	Length 746;
Best Local Similarity	-25.5%;	Pred. No. 1.8e-31;		
Matches 191;	Conservative 91;	Mismatches 290;	Indels 176;	Gaps 27;

```

Query Match 191: 10.1% Score 509; DB 5; Length 746;
Best Local Similarity 15.5% Pred. No. 1,8e-31;
Matches 191: Conservative 91; Mismatches 290; Indels 176; Gaps

Oy 344 LGNTY--TQKATFRVMAPISTQVNVLLYSATGS--VTKIIVPTASGHCWEMATV-NQN 398
Db 8 LGARIDAQAQAMLAERFVYSSRATREVELYKKNPTSGOVBARLALSKDPATQVSLSTPT 67
Oy 399 LENM-----YYMEVYVG-----GGS-----TRTAVDPATA 424
Db 68 IKNTYIGITGANYIYGRANGPMPPIDAAAMTKSATGATGVSDVDMAGNRPKLLDLPARE 127
Oy 425 IAPNGTGMIVD--LAKTDPAGNNSDKHI-----TPKNIDEVIYEM 464
Db 128 ISDPNPTACADGTYATGCAHRRKKDSGLCASKICALAUAATSVGSKPTBALDEVIEYV 187
Oy 465 DYRDESIDPNSGMMKKKYYLALTEKGT-KRPDNTKTIIDSLKOLGITHVQLMAYFAS--- 520
Db 188 HVR-----GLTRNDSDVPAAEERGTYYKGAARRAA---ALAALGAYAVEFLPVOEQND 236
Oy 521 -NSVDETPTODN-WGYDPRAVYDEPEQOYATNANGANARIKEFEMVLSLREHIGVMQV 578
Db 237 QNDVDNSTACDDNTWGYTTLVTFAPDRRIYADKSAAGPTREKMAVMKAFIDAGIKYIDV 296
Oy 579 VYNTT-----FAVDISPDFKIVPEYYIRMTQYIIPDOY-LEMKLXAERPV 625
Db 297 VYNTTGECPGMSGTDLGSVYVLLSFRGLDNPAYVSLSSDYKYPMDNTGCGVGNTRHPA 356
Oy 626 QKFIIDSLKYWNEXHIDGFEEDLMALGK-----DTMSKASEL---H 666
Db 357 QNLIVDSLATYKRDAIGVDGFEFLDASVIGNSCGHCGCFNPFKNDSGNALNIYAVELPRPA 416
Oy 667 AINPGIALYGERPWT-GCTSALPDDQLLTGKAQKG--MGVAFNDNLRNALDGNVPDSSA 722
Db 417 AGGAGADLIAEPMAIGGNS-----YQVGFPFAGNAENNGLYRALRKQNKLG 465
Oy 723 QGFATGATGLDAIKNGEGSINDFTSSPGTINYYVSHDNTYLTMDKIALSNRPDSF--- 779
Db 466 ETVTPTLATRPAGSNDLYG---DDGRKPWMSINPVAHAGHGFLLNDIYAINDQONQWP 522
Oy 780 -----ADRIKMDLAQAVVMTSGQVFMGGEGEMLRXKGGNDNSY 819
Db 523 YGPDGEGEDHNLSSNNGGIVALEQKRAARTGIALMLLSAGVPMITGDEALRTQGNNTY 582
Oy 820 NAGDAVNEFDMSRKAQYVDVENVYSGLIHLRLDHPAFR---MTTANFINSHLO----- 869
Db 583 NLDSSANMLYRSRSLLEADHETTYKRLIAFEKRAIPALRPANFYSASPTONNEMOULRFK 642
Oy 870 -----FLNSPEN-TVAAYELDHNKMKMGNIIVVYVPRKNTVATINLP---SGKMAI 916
Db 643 PDGAQADSATYNGADNALMARITDGSERGCDSASAIYVAINQWSCAVDFKLPMPETGQWY 702
Oy 917 NATSGKVESTLGAEGSVQVPGISMI 944
Db 703 RYTD---TATWNEGPNAVVALPGSETILI 726

RESULT 11
US-08-476-519-2
: Sequence 2, Application US/08476519
: Patent No. 5750876
: GENERAL INFORMATION:
: APPLICANT: Barry, Gerard F.
: APPLICANT: Kishore, Ganesh M.
: APPLICANT: Krohn, Bradley M.
: TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions

```



QY 579 YVNH-----FATQISDFDKIVPEYRTMIOVILPPTDQV-LEMKLAERPMV 625  
DB 328 YVNHTEGEGPMSGTDLGVNLSFRGLDNPAYYSLSDYKYPMDNTGVGNNTYTRHPIA 387  
QY 626 QKFTIDSLKYNVNEHYHIDGFRFDLMALGK-----DTRMSKASEL-----H 666  
DB 368 QNLIVDSLAYKRDALGVDGFRFDLASVLCNSCGHCFNFKNDGSMALNRIKVAELPRRPA 447  
QY 667 AINFGIALYGEPMW-GGTSALPDDQLLTKGAQK--MGAVFENDLRNALDGVNFDSSA 722  
DB 448 AGCAGALIEAPMAIGNS-----YVCGFPAGNAEMNGVLDALRKKNKLG 496  
QY 723 QCGFATGATGLDAIKNVEGSIINDFTSSPGFTINYVTSNDYTTIMDKIALSNPDS- 779  
DB 497 ETVYPTGLATREFAGSNDLYG---DDGRKPMHISINFEVAHDEFTINDLYAYNDKONNPPW 553  
QY 780 -----ADRIKMDLAQAVVMTSGVPMOGGEMLRKKGNDNSY 819  
DB 554 YGSPDGEDHNLMSNOGSIYAEORKAARTGLALLMLSAGVMTITGDEALRTOFONNNY 613  
QY 820 NAGDAVNEFMSRRAQYPDVFNYISGLIHLRLDHPAFR--MTANEINSHLO----- 869  
DB 614 NLDSAAWMLYMSRSLAEADHETTKRLIAFRKAPALRPANFYASDPTNGVMQLRMFK 673  
QY 870 -----FINSPEP-TVAEELTDHVNKKKMCNIIYYNPNKIVATINLP--SGKWI 916  
DB 674 PDGAODSAVFNAGDNLALMRIDGSEFDSASAIYVANGMSGAVDFKLPMTGTGRQWY 733  
QY 917 NATSGKVESTLQAECSVQYPCISMMI 944  
DB 734 RVTD-----TATWNEGPNAVVALPGSETLI 757

RESULT 13  
5457037-5  
; Patent No. 5457037  
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;  
; GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO  
; TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE  
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME  
; NUMBER OF SEQUENCES: 7  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/1,797  
; FILING DATE: 08-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 749,621  
; FILING DATE: 19-AUG-1991  
; APPLICATION NUMBER: 224,114  
; FILING DATE: 25-JUL-1988  
; SEQ ID NO:5:  
; LENGTH: 751  
5457037-5

Query Match 9.68; Score 487; DB 6; Length 751;  
Best Local Similarity 25.68; Pred. No. 9.7e-30; Indels 180; Gaps 23;  
Matches 186; Conservative 85; Mismatches 271;  
QY 344 LGNTY--TOKATTFKVAAPSTOVNVLLYDSATG-SVTKIIVPMASGHWAEATVNONL- 399  
DB 8 LGASYDAQAOANITFRVYSSQATRIVLVLSAGYGVQESATYITLSPAGSGVAAVTPVSSI 67  
QY 400 -----ENMYMEVTEGGS-----TRTAVDPYATAI 425  
DB 68 KAACTGAVVYGYRAMGPMPYASN-WKGSQAGFVSVDANGDRFNPMLLLDPYAOEV 126  
QY 426 A-----PNGTRGMIVDLA---KTDPAQWNSDKHIT-----PKNIEDEVYEMD 465  
DB 127 SQDPLNSNONGNVEFASGASVYRTDSGIYAPKGVVLPVSTQSTGKPTRAQKDDVYIEVH 186  
QY 466 VRDES-IDPNSGMKMKGYLALTEKGTGKGPNDVNTGIDSLKQLGITHVOLMPV-----FA 519

DB 187 VRGETDQSTIPAOYRGYTAGLKAS-----YLASLGVTAVEPLPVOETONDA 235  
QY 520 SNSVDETDPTDQNMNGYDPRNVDVEGOYATNANGNARIKEFEKVELSLHREHIGVMDYV 579  
DB 236 NDVVPNSDANQNYGMVMTENEFSPDRRYAYNKAAGPTAEQAVQAFHNHNGIVYMDYV 295  
QY 580 YNH-----TFATOISDFDKIYPEYRTMIOVILPPTDQVLEMKLAERPMVQK 627  
DB 296 YNHAEGETWTSODPTATITSMRGLDNTTYIELTSGNQYFYDWTGICANENTYNTVAQN 355  
QY 628 FIIDSLKYNVNEHYHIDGFRFDLMALGKDTMS-----KASELAIN- 669  
DB 356 LIVDSLAYWANTMGVDGFRFDLASVLCNSCLNGAYTASAPCPGNGVNFDAADSNVALNR 415  
QY 670 -----PGIALYGEPMW-GGTSALPDDQL-LTRGAQKMGVAVFENDLRNAL 713  
DB 416 ILRETVYRPAAGSGGLDLEAFEPMAIGNSY---QLGFGPGWEMNG--LFRDRLR--- 466  
QY 714 DGNVFDSSAQCFAATGATGLDAIKNVEGSIINDFTS--SPGFTINYVTSNDYTTIMDKI 770  
DB 467 -----QAOHELGSMTIYVTDANDSESSNLFQSSGGRSPWNSINFIDVHGDTLADYV 519  
QY 771 ALSNPDS-DRIMKMDLAQAVVMTSGVPM 802  
DB 520 SCNANNSQAMPYGPDSGISTNYSWDOGMSAGTGAADVORRAARTGMAFEMLSAGTFLM 579  
QY 803 QGGEEMLRKKGNDNSYNAGDAVNEFMSRRAQYPDVFNYISGLIHLRLDHPAFR----- 857  
DB 580 QGGDEVLETTQCCNNNAVYLLDSSANMLTYSMTDQSNFETFAORLIAFRKAPALRPSSWY 639  
QY 858 ---MTANEINSHLOFLNSPNTYAVELTDHVNKKDKC---NIIYVNPKNVTATINLP 911  
DB 640 SSQLTWYQPSGAVAADSVMYNTSNTYALVAINGPSLSDSNSIYVANGMSVTFLLPA 699

RESULT 14  
5457037-3  
; Patent No. 5457037  
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;  
; GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO  
; TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE  
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME  
; NUMBER OF SEQUENCES: 7  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/1,797  
; FILING DATE: 08-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 749,621  
; FILING DATE: 19-AUG-1991  
; APPLICATION NUMBER: 224,114  
; FILING DATE: 25-JUL-1988  
; SEQ ID NO:3:  
; LENGTH: 750  
5457037-3

Query Match 9.58; Score 482; DB 6; Length 750;  
Best Local Similarity 25.48; Pred. No. 2.4e-29;  
Matches 183; Conservative 86; Mismatches 271; Indels 180; Gaps 23;  
QY 344 LGNTY--TOKATTFKVAAPSTOVNVLLYDSATG-SVTKIIVPMASGHWAEATVNONL- 399  
DB 7 LGASYDAQAOANITFRVYSSQATRIVLVLSAGYGVQESATYITLSPAGSGVAAVTPVSSI 66  
QY 400 -----ENMYMEVTEGGS-----TRTAVDPYATAI 425  
DB 67 KAACTGAVVYGYRAMGPMPYASN-WKGSQAGFVSVDANGDRFNPMLLLDPYAOEV 125  
QY 426 A-----PNGTRGMIVDLA---KTDPAQWNSDKHIT-----PKNIEDEVYEMD 465  
DB 126 SQDPLNSNONGNVEFASGASVYRTDSGIYAPKGVVLPVSTQSTGKPTRAQKDDVYIEVH 185  
QY 466 VRDES-IDPNSGMKMKGYLALTEKGTGKGPNDVNTGIDSLKQLGITHVOLMPV-----FA 519



```

Db 186 VRGFTEDTSTIPAOYRGTYAGAGLAKS-----YLSLGVTAVEFLVQSTQND 234
QY 520 SNSVDETPDODNMGYDRNTDVEGOYATNANGNARIKEFEKEMLSJHREHIGVNDVY 579
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 NDVPNSNANQNYQMYKMENTFSPDRRYAYNKAAGPTAEQAOAHNAGIKYMDVY 294
QY 580 YNH-----TFATQISDFDKIVPEXYRTMIOVILPTDOVLEMKLAEPRMYOK 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 YNHAEQGTWTSDEPTATYISMRGLDNTTYEILTSNQYFYDNTGIGANFNTYMTVAQN 354
QY 628 FIISLKWVWEYHIDGRFDLMALGKDTMS-----KAASELHAIN- 669
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 LIVSLAWMTMAGVDGFRFDLASVLSCLMGATASAPNCPNGCYFEDADSDNSVAINR 414
QY 670 -----PGIALYGEPTW- GGTSALEPDDQ- LTKGAKQMGVAFENDILRNAL 713
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 415 ILRETVYPPAAGSGGLDFAERMAIGNSY-----OLGGFPQGSSEWNG- LFRDSIR- 465
QY 714 DGNVFDSSAQGFATGATGLTDAIKNGVEGSIINDFTS---SPGETINYVTSHDNYTLMDKI 770
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 466 -----OAOEHLGSMITVYQDANDFGSSNLFQSSGRSPWNSINFLDHDGMTLKDY 518
QY 771 ALSNPDEA-----DRIMDELAQAVVTFSGVPPM 802
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 519 SCNGNANSQAMPYGPDSGDTSTNSWDMQMSAGTGAAYNORRAARTGMAFEMLSAGTPLM 578
QY 803 OGGEEMLXKXGNDNSYAGDAVNEFDMRKAQYPDVFNYSGLIHLRLDHPAFR----- 857
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 579 OGGDEYLTLOCNNAYNLDSSANMLTYSWTDOSNFTYFORLAFRAHAPALRPSMY 638
QY 858 ---MTANEINSHLOFLSPENTVAEYELTDHVNKDKMG---NIIVYVNPNTKVAATINLPS 911
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 639 SGSQLTWYQPSGAVADSVNWNMTSNYAIAYAINCPLGDSNSIYVAYNGWSSSVFTFLPA 698

RESULT 15
US-09-129-075-4
: Sequence 4, Application US/09129075
: Patent No. 6087149
: GENERAL INFORMATION:
: APPLICANT: Tsutsumi, No. 6087149iko
: APPLICANT: Bisgard-Frantzen, Henrik
: APPLICANT: Svendsen, Allan
: TITLE OF INVENTION: Starch Conversion Process
: FILE REFERENCE: 5262.204-US
: CURRENT APPLICATION NUMBER: US/09/129, 075
: EARLIER FILING DATE: 1998-08-04
: EARLIER APPLICATION NUMBER: 0787/97
: EARLIER FILING DATE: 1997-07-02
: EARLIER APPLICATION NUMBER: 60/055, 867
: EARLIER FILING DATE: 1997-08-13
: EARLIER APPLICATION NUMBER: PCT/DK98/00304
: EARLIER FILING DATE: 1998-07-02
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 726
: TYPE: PRT
: ORGANISM: Rhodothermus marinus
US-09-129-075-4

```

Query Match 9.4%; Score 475; DB 3: Length 726;  
 Best Local Similarity 26.8%; Pred. No. 8e-29;  
 Matches 166; Conservative 88; Mismatches 211; Indels 154; Gaps 25;

```

QY 344 LGNTYOKATTFKVAAPTSTOVNLLYDSANG-SYTKIIVPMTASGHGWAEATVNONLENN 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 22 LGATWDGAGVNFALYSQAAEAVELFDHPDPAPSRTIEVETRGPIWHVYLPGLRPGO 81
QY 403 YMYEVYVQO-----GSTRTANVPYATAA-PNGTRGMIVDLAKTDPAG----- 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 82 LYGYRYGPRPEEGHREPNKVLDPYAKAIGRLRWHDLSFGYKIGDPAGDLSFEED 141
QY 445 -----WNSDKHITPKNI--EDEVIEYMDVRFDS-IDPNSGMKNGKYLA 485
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 SAPYAPLAGAVNGCFEWMGDR---PPRIWEDTIIIEYHVKGIITKLHPEVPEPLKRTYIG 198
QY 486 LTEKGTGPDNKTGIDSLKQIGITHVQMLPVFAS---NSVDETPDODNMGYDRNRD 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 LF-----CEPVLEHLKQLGVTTIQLLPYHAKVHDBHLVER--GLNNWGYNPLCYE 247
QY 542 VPEGYATNANGNARIKEFEKEMLSJHREHIGVNDVYNHTEFATQI-----SDFOKIVE 597
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 APEPEYATNGPISA-VREFKMYRALHAAGFEYIVDVVYNNHGGEGVGLPTLSFRGIDNR 306
QY 598 YYYRT-----MIQVILPTDOVLEMKLAEPRMYOKFIISLKWVWEYHIDGRFDLMA 651
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 AYKADPNRPRLVDYTGNTLDV-----GNRYVLIQDPSLVQYKLIAPMDVGP- 362
QY 652 LIGKDTMSKAASELHAIN-----PGIALYGEPTWGTGTSALPDDQLITTKGA 696
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 AL-----ARELYDVDMLSFEFGVYIQDDPVLQYKLIAPMDVGP- 404
QY 697 QKG---MGVAFENDILRNALDGNVFDSSAQGFATGATGLTDAIKNGVEGSIINDFTS--- 750
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 405 QVGHFPWQTEWNGRYRDV-----RRFRGDRGNGEFAFATPAGSSDLYERSGR 455
QY 751 PGETINYVTSHDNYTLMDKIALS---NPNDSADRIKMD- 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 PRASINFTYAHGFTLEDVYSTTKKNEANLEGNRCGMENSTNGCVGPTODPSYLAC 515
QY 788 ---LAQAVVYTF---SGGVFPQOGGEEMLXKXGNDNSYAGDAVNEFDW---SRKAQYPD 838
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 516 RALKRSLISTFLFSQGVPMILGDELSTQIGNNNAACODNEISYNNQLODTRKQOF-- 573
QY 839 VFNYVSGLIHLRLDHPAFR 857
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 574 -LEFVRQTIWKRQHSFR 591

```

Search completed: April 22, 2001, 09:05:43  
 Job time: 4027 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 08:57:44 ; Search time 51.59 Seconds

(without alignments)  
47.377 Million cell updates/sec

Title: US-09-262-126C-2

Perfect score: 5060  
Sequence: 1 MAKKLIYCLSVCLVTWAF.....VPGISMILHOFVSPDHGKK 956

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 13044 seqs, 2556675 residues

Total number of hits satisfying chosen parameters: 13044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New: \*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5054	99.9	956	US-09-262-126C-2	Sequence 2, Appl
2	1039.5	20.5	718	US-09-262-126C-3	Sequence 3, Appl
3	787.5	15.6	1091	US-09-262-126C-4	Sequence 4, Appl
4	528.5	10.4	793	US-09-463-238-5	Sequence 5, Appl
5	506	10.0	766	US-09-463-238-6	Sequence 6, Appl
6	504	10.0	838	US-09-463-238-14	Sequence 14, Appl
7	293.5	5.8	397	US-09-463-238-15	Sequence 15, Appl
8	270	5.3	878	US-09-463-238-4	Sequence 4, Appl
9	269	5.3	883	US-09-463-238-19	Sequence 19, Appl
10	265	5.2	379	US-09-739-449-8397	Sequence 19, Appl
11	196	3.9	1833	US-08-945-567C-4	Sequence 8397, Ap
12	196	3.9	1992	US-08-945-567C-3	Sequence 4, Appl
13	165	3.3	530	US-09-739-449-12578	Sequence 3, Appl
14	161.5	3.2	533	US-09-739-449-9531	Sequence 9531, A
15	150	3.0	533	US-09-739-449-9662	Sequence 9662, Ap
16	149	2.9	735	US-09-739-449-8401	Sequence 8401, Ap
17	146	2.9	478	US-09-327-563B-10	Sequence 10, Appl
18	124.5	2.5	725	US-09-739-449-12023	Sequence 12023, A
19	123.5	2.4	1976	US-09-739-449-9987	Sequence 9987, Ap
20	123	2.4	549	US-09-327-563B-6	Sequence 6, Appl
21	119.5	2.3	485	US-09-327-563B-12	Sequence 12, Appl
22	116.5	2.3	514	US-09-327-563B-4	Sequence 4, Appl
23	116	2.3	868	US-09-739-449-12606	Sequence 12606, A
24	115.5	2.3	1205	US-09-806-080-4	Sequence 4, Appl
25	115	2.3	472	US-09-739-449-11574	Sequence 11574, A
26	114.5	2.3	512	US-09-327-563B-2	Sequence 2, Appl
27	113.5	2.2	483	US-09-327-563B-13	Sequence 13, Appl

28	107.5	2.1	1234	US-09-739-449-8719	Sequence 8719, Ap
29	106.5	2.1	4563	US-09-802-640-32	Sequence 32, Appl
30	104.5	2.1	310	US-09-739-449-9428	Sequence 9428, Ap
31	104.5	2.1	485	US-09-739-449-11070	Sequence 11070, A
32	103.5	2.0	637	US-09-739-449-9991	Sequence 9991, Ap
33	102	2.0	647	US-09-739-449-12524	Sequence 12524, A
34	101.5	2.0	1471	US-09-421-124-188	Sequence 188, App
35	101.5	2.0	1471	US-09-421-124-188	Sequence 188, App
36	100	2.0	703	US-09-739-449-9778	Sequence 9778, Ap
37	100	2.0	1494	US-09-421-124-186	Sequence 186, App
38	100	2.0	1494	US-09-421-124-186	Sequence 186, App
39	99.5	2.0	793	US-09-739-449-11850	Sequence 11850, A
40	98	1.9	697	US-09-739-449-13223	Sequence 13223, A
41	96.5	1.9	943	US-09-193-562D-32	Sequence 32, Appl
42	96	1.9	665	US-09-739-449-12462	Sequence 12462, A
43	96	1.9	980	US-09-628-359-11	Sequence 980, Ap
44	96	1.9	996	US-09-739-449-9909	Sequence 9909, Ap
45	95.5	1.9	551	US-09-739-449-11423	Sequence 11423, A

ALIGNMENTS

RESULT 1					
US-09-262-126C-2					
Sequence 2, Application US/09262126C					
GENERAL INFORMATION:					
APPLICANT: Shetty, Jayarama K.					
TITLE OF INVENTION: Modified Forms of Pullulanase					
FILE REFERENCE: GC396-2					
CURRENT APPLICATION NUMBER: US/09/262,126C					
CURRENT FILING DATE: 1999-03-03					
NUMBER OF SEQ ID NOS: 9					
SOFTWARE: FastSeq for Windows Version 3.0					
SEQ ID NO 2					
LENGTH: 956					
TYPE: PR					
ORGANISM: Bacillus deramificans					
FEATURE:					
NAME/KEY: VARIANT					
LOCATION: (1)...(956)					
OTHER INFORMATION: Xaa - Any Amino Acid					
US-09-262-126C-2					
Query Match					
Best Local Similarity 99.9%; Score 5054; DB 5; Length 956;					
Matches 956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MAKKLIYCLSVCLVTWAFNFKGSAHADGNTTIIHYFCFAGDPOMSLMMKPRDGG	60		
DB	1	MAKKLIYCLSVCLVTWAFNFKGSAHADGNTTIIHYFCFAGDPOMSLMMKPRDGG	60		
QY	61	GAEYDFNPGADSGAASADIPGNSQVGIIVRTQDMKDVSAADRYIDLKSGNEVLLVYG	120		
DB	61	GAEYDFNPGADSGAASADIPGNSQVGIIVRTQDMKDVSAADRYIDLKSGNEVLLVYG	120		
QY	121	NSQIFPNKDEAEPAAFPAYSNATLDSNOVLKLSQPLTIGEXSGFTYHDDPANKDIPY	180		
DB	121	NSQIFPNKDEAEPAAFPAYSNATLDSNOVLKLSQPLTIGEXSGFTYHDDPANKDIPY	180		
QY	181	TSVKDASLQDQYAVYLAGTFQHFQFGSDNAPDHSTLLKVTNNLQFSGDLPEGNYYOK	240		
DB	181	TSVKDASLQDQYAVYLAGTFQHFQFGSDNAPDHSTLLKVTNNLQFSGDLPEGNYYOK	240		
QY	241	VALNDSMNNSYPSDNLNLTPVACGAHVTFYIPSTHAVYDTINNPNADLOVESGVTDLY	300		
DB	241	VALNDSMNNSYPSDNLNLTPVACGAHVTFYIPSTHAVYDTINNPNADLOVESGVTDLY	300		
QY	301	TYTLGDDPVSHSLSTOTGQYQAKOYIPRNVLNSQYYSGGDLGNTYQKATTFVWAP	360		
DB	301	TYTLGDDPVSHSLSTOTGQYQAKOYIPRNVLNSQYYSGGDLGNTYQKATTFVWAP	360		

```

QY 361 TSTOVNVLVLSBATSQSVKIPYMTASGHWBETATONLENNYVAYEATGOGSTETADP 420
Db 361 TSTOVNVLVLSBATSQSVKIPYMTASGHWBETATONLENNYVAYEATGOGSTETADP 420
QY 421 YATAIAPNGTRGMIYDLAKTDPAGANSDBKHITPKNIEDEVITYEMDVRDPSIDPNSGMKK 480
Db 421 YATAIAPNGTRGMIYDLAKTDPAGANSDBKHITPKNIEDEVITYEMDVRDPSIDPNSGMKK 480
QY 481 GXYLATFEKGRKGPBNVTCIDSLKOLDITTHOQLMPVASNSVDDETDPQDMWGDPDRY 540
Db 481 GXYLATFEKGRKGPBNVTCIDSLKOLDITTHOQLMPVASNSVDDETDPQDMWGDPDRY 540
QY 541 DVEPEGVATNANGNARIKEFEKEMVLSLHREHIGVMDVYNTTFATQISDFKIVPEYXY 600
Db 541 DVEPEGVATNANGNARIKEFEKEMVLSLHREHIGVMDVYNTTFATQISDFKIVPEYXY 600
QY 601 RTMIOVIITPDQVYLEMKIAXERPVOKRTIDSLKATWNEYHIDGRFDLMLLKGDTMSK 660
Db 601 RTMIOVIITPDQVYLEMKIAXERPVOKRTIDSLKATWNEYHIDGRFDLMLLKGDTMSK 660
QY 661 AASELHAINPGIAXGEPWTCSTALPPDOLITKAQGMGVAAYNDNLRLNALDGNVDS 720
Db 661 AASELHAINPGIAXGEPWTCSTALPPDOLITKAQGMGVAAYNDNLRLNALDGNVDS 720
QY 721 SAQGFATGATGLTDAIKKGVBGSIINDFTSSPEGTINYTSHDNYTLMDKIALSNPDS 780
Db 721 SAQGFATGATGLTDAIKKGVBGSIINDFTSSPEGTINYTSHDNYTLMDKIALSNPDS 780
QY 781 DRIKDELAQAVMVSQGVPEFMQGEEMLRKKGNDNSYNAGDAVNEPDMRSKAQYPOVF 840
Db 781 DRIKDELAQAVMVSQGVPEFMQGEEMLRKKGNDNSYNAGDAVNEPDMRSKAQYPOVF 840
QY 841 NYISGLHLRLDHPAFRMTTANEINSHLOPLNSPENTYAYELTDHVYNDKDKNIIIVYNP 900
Db 841 NYISGLHLRLDHPAFRMTTANEINSHLOPLNSPENTYAYELTDHVYNDKDKNIIIVYNP 900
QY 901 NKTVATINLPSCKMAINATSGKVGCESTLGOAGSVQVPGISMMILHOEYSPDHGK 956
Db 901 NKTVATINLPSCKMAINATSGKVGCESTLGOAGSVQVPGISMMILHOEYSPDHGK 956

RESULT 2
: DS-09-262-126C-3
: Sequence 3, Application US/09262126C
: GENERAL INFORMATION:
: APPLICANT: Miller, Brian S.
: APPLICANT: Shetty, Jayarama K.
: TITLE OF INVENTION: Modified Forms of Pululanase
: FILE REFERENCE: GC396-2
: CURRENT APPLICATION NUMBER: US/09/262.126C
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 718
: TYPE: PRY
: ORGANISM: Bacillus subtilis
: US-09-262-126C-3

```

[illegible]

Qy	354	TEKWAPESTOVNVLXDYSAGSVTKIYVPMASGHWEATVNONLENNWYMEVTOGS	41.3
Db	116	VFKWAPARATSAAKL--SHRNSGRFTQMRLKGYVATVYGDLLGLEYLPTCINSE	173
Qy	414	TRTAVDPATALAPNGTRGMTVDLAKTDPAGCNSDKHITP-KNIEDEYIYEMUDROFSID	472
Db	174	WMETVDOYAKAVLVTVNGEGVVL---RPMQKMTAP-LKPFHSHPVDVAIVYETHLRFSIH	228
Qy	473	PNSGKNNGKGLATLETKGKPRDVKVNGIDSLKOLGTHVOLMPVAFNSVSDIEDDPQD-	531
Db	229	ENSGINNGKGLATLETDPTQVANSSSSGLAYVKLGTVHVELLPVNDPAGVDEEKLIDAY	268
Qy	532	NMGIDPRIVDYREQYATNANG-NARIKEREWYLSLREHIGVYNDVYNHTFATQISD	590
Db	289	NMGVNPRLHFEPPESSYASNPDPQTRKTELQMINTLHQHGLRVLIDVFNHYKRENSP	348
Qy	591	FDKJIEPEYVYRTMIOVLIPDOV-LEMKLAKERPVOVKFIISLSKYVNEHIDGPFEDL	649
Db	349	FEKTYPGVFFR-HDECGSPNSGTGVNDIASERSMARKFIADCYVYVLEXYNDGPFEDL	407
Qy	650	MALLGKDTMSKAASELHAINGCLALYGERPWTGTSALRPDOLLTKGAKQMGVAVFNDNL	709
Db	408	LGILIDIDIVLWKKRATKAKGKILFFEGMDLAPRLRHEKKAALANAPRMGQICOFFNDMF	467
Qy	710	RNALDGNVFDSSAOGFATGATGLDALKNGVEGS----INDETSSPGETINVTSHDY	764
Db	468	RDVAGNTPFLHAKTGAFALNGCESAOAVMHGIAGSGKALAPIVBERSQSINYESHDH	527
Qy	765	TLMPKIALSNDEBADIKMDELQAVVMVMSQGVPRMOGGEELRYKKGNDNSYNAGDA	824
Db	528	TFMDKMSALFOENDSKRRKORLAVAILLAQVAPRIHSGQEEFRKQOEVENSYSOSDS	587
Qy	825	VNEFDMSKRAQYRDVFNYSGLIHLRLDHPAFRMTJANEINSLQFTNSBPNTVAYELTD	884
Db	588	INOLDMDRRETFKBDVNHIRRLISLRANHPAFRLRSADAIORNHIECTLTLEHLAYELXD	647
Qy	885	HVNCKMKNITLVVYNPNPTVATVIMPSG--KMAINTNSGKVGESTLQAGSVOYOPISM	942
Db	648	LDEVDKMDIIVIHNASDVSFEMRLPNDIPRLCLDPSGQOEDEPT--EIKTYAVNKGIGT	705
Qy	943	MILH 946	
Db	706	VILY 709	

```

RESULT      3
US-09-262-126C-4
: Sequence 4, Application US/09262126C
: GENERAL INFORMATION:
: APPLICANT: Miller, Brian S.
: APPLICANT: Shetty, Jayarama K.
: TITLE OF INVENTION: Modified Forms of Pullulanase
: FILE REFERENCE: GC996-2
: CURRENT APPLICATION NUMBER: US/09/262.126C
: CURRENT FILING DATE: 1999-03-03
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 1091
: TYPE: PRT
: ORGANISM: Klebsiella pneumoniae
US-09-262-126C-4

Query Match          15.6%; Score 787.5; DB 5; Length 1091;
Best Local Similarity 25.3%; Pred. No. 6,5e+48;
Matches 288; Conservative 154; Mismatches 365; Indels 327; Gaps 48;

OY      22 VKGSAHADGNTTTIIVYFCPAG-----DYQPWSLMMWPKKDGCAEYDENCOPADSPG 74
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db      50 VPGENVAVENQA--VHILVDIAGITSSAADYSKKNLITLMN-----NEHCDAIS 97
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|

YY      75 AVASADIDGNPSQVGIIIVRTODMTKDV-----ADRYTIDLKSGNEWVLVEGNSQILEYNEK 129

```

```

Db 98 APVA-----DW-NDVSTPPSGSDKY-----GPWVPLNKE-----127
QY 130 DAEDANPAVSNAYLDASNOYLVKLSOPLTLEGXSGFTVHDDTANKDIPYTSVKDASIG 189
Db 128 -----SGCINIVYRSDTKLISDLRVAFCD-----FT--DRYTSVLAGNSAVYDSR-- 172
QY 190 QDVTAVLAGFQHIIFG-----GSDMAPDNHSTLL-----KVTNNLYQFSGDLP--ECN 236
Db 173 -----ADAFRAAFVGAALAEHAWDKN--TLLWPGGQDKIVRLYSHSSKVAADGEGK 223
QY 237 YQYKVALINDSNNSYP---SPINILTVPAGAHYFSTIPESTHAYDIDINPNAD--IQY 291
Db 224 F-----TDRLKLPPTTVSOOVSNRFP-----HL-----SSYAFKLPDANAVDELLOG 267
QY 292 ESGVKTDLVYTLGEDPVSHTLSIQTDGYQAKQYIPRNVLNNSOYYSGDLP--GNTYT 349
Db 268 ET-----VAIAAEDGLISATQVQTAG-----VLDDA-YAAAEALSTGAOLA 310
QY 350 QKATFPKVAWPTSTOVNVLVYDSATGVTAKIVPMTA--SGHGVEATVQNLENNYMYEV 408
Db 311 DGVTFRVWAPTAAQOVVVYV--SADKVIYIGSHPMTRDSASGMSMGSGSDLKGAFYRYAM 369
QY 409 TGGOSTRAV-----DPYATAIAPNGTRGMIVDL--AKTOPAGNS--DKHITPKNI 456
Db 370 TVYHPOSRRKQOYEVTDFYASHLSTNSEYSOVVDLNDLALPDGMDNLJMPHAKTKRADL 429
QY 457 EDEVIYEMDVDFST--DPNSGMKNKGYLALTEKTRKPDVNGTIDSLKOLGITHVOLM 515
Db 430 AKMTIHESHIDLSAMDQTVPAELRGKYLAL--TAGDSNWOHLKTLISASGVTHVELL 485
QY 516 PVFASNSVDE--TDPTQD-----NAGYPRNRYVPEGQVATNANGARIKEFEKMYLSLREHIG 531
Db 486 PVFDLATYNERSDKYADIQOFPFSLCEVNSAVKSESEFAGCDSSTVEVLNQLKQSDQ 545
QY 532 -----NAGYPRNRYVPEGQVATNANGARIKEFEKMYLSLREHIG 573
Db 546 DNPVOVALNTLVAQTDSTYMWGYDPEHYTVPEGSTATDEGTRIKEFRTMIOAI--KODIG 604
QY 574 VN--MDVYNNHTFAATQISD--FDKIVPEXYR-----TMIVYIITPDVLEKMLAE 621
Db 605 MNVIMDVYNNHTNAGPDRTSYVDKIVPWYQRLNETTGSVESATCCSDSPERHMRFA-- 663
QY 622 RPNWOKFIIDSLKYWNEYHIDGFERDLMALLGKDTMSKASELHAINPGIALYGEPTWG 681
Db 664 -----KLIAJSLAVWTTDYKIDGFERDLMGHPKQIILAMERIKALNPDIYFGEQWDS 718
QY 682 CTSALPDDQILTKGAKQKMGVAVFENDNLNML--DGNVFDSD--SAGGFATGA----- 729
Db 719 NOS--DREILASQILNKGTIGTSDRLRDSVRCGGPDSGDALNQNOGIGSAGAVLPNE 776
QY 730 -----TGLTDAIKKNGEVSINDF-----TSSGETIN 756
Db 777 LASLSDQVRLADLTRLGMAKNLADFYAIDKGAAGKSEIDYNGARGVYAADTEVYN 836
QY 757 VYTSNDNTYLMDKIALSNPNDS--ADRIKMDLQAVVWTSQGVPPMOGEGEMLRXKGN 815
Db 837 YVSKHNOJTLDMISYSKASEADLATRVQMQVSLATVWLGQIAFPDQGSSELLKSKPT 896
QY 816 DNSTYAGAVNEFDMS-----RKQYVDVFN 842
Db 897 RDSYDSGMPNRYVDSIADNNNYNMGMPRI SDGSNVEYITRVKENVATPGEAELOMRAF 956
QY 843 YSGILHRLDHPAFRMTTANEINSHLOFLNSPENTVAYELTDHVN-----KDKMGN 893
Db 957 YQELTELKSSGLPTLGGASAMKRVDFRNTGSDQOAGLIVTVVDGMAKAGASLDSRLDG 1016
QY 894 IIVVYNPKNVATINLPSCG--MAINATSGKVGESTLGO-----AEGSOVQPGISMMIL 945
Db 1017 LVVAINAAPESTRINEFAGETLQLSAIOOTAGENSLANGVOIAADGTVTLPLAMSVAVL 1074

```

RESULT 4

```

US-09-463-238-5
; Sequence 5, Application US/09463238
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Allison M
; APPLICANT: Bustos Guillen, Regia
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463.238
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-5

Query Match          10.4%; Score 528.5; DB 5; Length 793;
Best Local Similarity 26.4%; Pred. No. 5.1e-30;
Matches 187; Conservative 100; Mismatches 255; Indels 165; Gaps 29;

QY 340 SGDDL--GNTYQAKATTEKVAWPTSTOVN--VLYDSATGVTYKIV--PMTASGHGW 391
Db 80 SGKPLPFGATATADGGVNVAVFESRNATATLITLSDLPKEKVTBQIFLDPLANKTGDW 139
QY 392 EATVQNLENNYMYEYVGO-----GSRRTAVDPATAIANGTRGMI----- 434
Db 140 HVFLGDEFENMLGYKFKDCKFCBEGHFDSSQIYVLDPAKAVSRGEYGLVGPEDDQCP 199
QY 435 ---VDLAKTDPAGNNSDKHIT--PRNIEDEVILEMDVRDFSIDPNSGMKNKGYLALTEK 489
Db 200 PMAGVWPSASDQFDMEDGLPLKFPQ--RDLVIEYKHNVRGFTNHESEETKPYCTYLGVYEK 257
QY 490 GTKGDNWKTKGIDSLKOLGITHVOLMPVAFASVDE-----TPTQDNMGIDPRANDV 542
Db 258 -----LDHLKEIVNCIELMPCHEFNELEYYSYNSVLYGKFFNFWGYSTVNFPS 306
QY 543 PEGQVATNANGN--ARIKEFEKMYLSLREHIGVMDVYNNHTATQ--ISDFDKIV 595
Db 307 PMGRISAGLSKCGIGALINEFYLKVEAKRGCEVIMDVYNNHTAEGNENCPILSFRGID 366
QY 596 PEYRYRTMIQVLIPTDOYLEMK-----LKAERPWQKFIIDSLKYWNEYHIDGFERDLM 650
Db 367 NSVF-----TLAPGEFYNVSGCGTFCNPNPIYRQFIVDCLRWVTEHMDVGFEDLA 421
QY 651 ALLGDTMSKASELHAINP--GIALYGEPTWGCT-----SALPDDQLTKGA----- 696
Db 422 STL-----TRSSSNWNAVNVGNSIDGVITTTPLTSPPLIDMSINDPIL--RGVKLAE 475
QY 697 -----OKG-----GV--AVFENDNLNMLDGNVFDSSAQGFATGTLTDAIKKNGEVS 743
Db 476 AMDCGGLYOVGFPHMGIMSEWNGKRYDNV-----RQFTIKGDGSGAPAECLGGS 526
QY 744 INDPTS--SPGETINVTYSHDNTYLMDKIALSNPNDSADRIKMD----- 786
Db 527 PMLYQKGRKPKMNSINFVCAHDGFTLADLVYTNKNKINLANGDKNKGHNHNSMNGGEG 586
QY 787 ELAQAVY-----MTSOGVPPMOGEGEMLRXKGNDSYVAGDAVNEFDMS 831
Db 587 EPASIFVKRLRKROMKNFELCLMVSOGVPMIYMGDEYGHTKGGNNNTYCHDNYINVFMD 646
QY 832 RKAQYVDVFNYSGLI-----HLRLDHPAFRMTTANEINSHLOFLNSPE-----NTV 878
Db 647 KDESSSDPLRRCGLMTKFRHECESLGLD-----GFTTAERLQWNGHGTPTPDMSSTRIV 702

```

QY 879 AVELTDHVNKDKMGNIIVVYNPNKTVATINLPSC---KWAINATSGK 922  
| : | | | : : : | | | : |  
Db 703 AFTLVDRKVR---GELYIAFNASHLPVTITLPDRPGYRQDFVDTGK 745

RESULT 5  
 US-09-463-238-6  
 : Sequence 6, Application US/09463238  
 : GENERAL INFORMATION:  
 : APPLICANT: Edwards, Elizabeth A  
 : APPLICANT: Smith, Allison M  
 : APPLICANT: Bustos Guillen, Regia  
 : APPLICANT: Martin, Catherine R  
 : APPLICANT: Plant Bioscience Limited  
 : TITLE OF INVENTION: Starch Debranching Enzymes  
 : FILE REFERENCE: 97.118  
 : CURRENT APPLICATION NUMBER: US/09/463.238  
 : PRIOR FILING DATE: 2000-01-21  
 : CURRENT APPLICATION NUMBER: PCT/GB98/02280  
 : PRIOR FILING DATE: 1998-07-30  
 : PRIOR APPLICATION NUMBER: GB 9716185.5  
 : PRIOR FILING DATE: 1997-07-31  
 : NUMBER OF SEQ ID NOS: 30  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 6  
 : LENGTH: 766  
 : TYPE: PRT  
 : ORGANISM: Solanum tuberosum  
 : US-09-463-238-6

Query Match	10.0%	Score 506	DB 5	Length 766
Best Local Similarity	26.2%	Pred. No. 1.8e-28		
Matches 172	Conservative 93	Mismatches 247	Indels 144	Gaps 27

QY	361	TSTOVNLLYSATSGVYKIVPMASGH-GVMEWTV-----NONLENNY-----	4093
Db	103	TESGINFATFESQIHSAVTLCITLPEKSVYHDSIELALDPOKNRGDIYHICIKELPOGCVL	
QY	404	YMEVTCOGGSTRFA-----VDPYATAIAPNGTSG-----MIVDLAATD-----PAGW	44.5
Db	163	YGYRIDGPRNMHECHREFDSDIILVDPRAKILLEGRRVGDSDSNKMCFFGTGYDFNSLPEDW	22.22
QY	446	NSDHRIPRKNIEDVYIEMOYRDSDIDPNSGM--KNKGKIALTEKTKGPDNWKIDS	50.3
Db	223	GENKILPRIPKDLVYTEMNRAFTAEOTSLDODQSGYGLIEK-----Iph	27.1
QY	504	LKOLGITHVOLMPAFASNSVD-----ETDPTODNMGEPRNRYVPEGOATNANGARI	55.57
Db	272	LLEIGVNAVELLRFVEFDELELQRRPMPROIMTWGYSITINFPAPRSRASVCGGGYVRA	33.11
QY	558	K-EFEKAWLSIAREHIGVNMDDVYNNH-----FATQISDFKIVPEYYTYRTMIOYII	60.8
Db	332	SWERKEWYKALHGAIEVILIDVYNNHNEADDEBNRYTTSRGIDNKV--YUWVDLN--	38.95
QY	609	PTDGVLEK-----FLXAEPRVAKFIIDSLKYWYNEZHIGFREDLALLGKDT-----	65.7
Db	386	NNAQLNPAGCGSNFCNCHNPTVMMLIESLRHWYTEHVHDSFRDLASVLCRGIDGPRIN	44.5
QY	658	---MSKAASELHAINPQIALYGERWYTG---TSALPDDOLLTKGAOKGAGVAFENDNLR	71.0
Db	446	APPLVAKISK-DVSLSRCKIIAEPMDCGGIYLKGFPMWRMAEMWNG-----YRDDIR	49.8
QY	711	NALDGNFVDSAGQ-FATGATGCLDIAIKNGEGSINFOTSSPGCTINUYVSHNUTYTLMDK	76.9
Db	499	RFING---DAGMKCNFATRLAGSADLYR-----VN--KKRPYHSVAFVTAHHGFTLYDL	54.7
QY	770	IALSNP-NDSEAD-----RIKMDLAQAVVTSQGV	80.0
Db	548	VSYNNKHNDANGEGCNDGNDNFSGNCIEGTSDANINALRSROMKFNHIALAVSGTDP	60.7
QY	801	FMOGGEMLAKKGGNDNSYANGAVNFEBSR--KAQYPDVFNYYSGLIHLRLDHPAFRMT	85.9

Db 608 MMLMGDEGHTRYGNNNSYGDITLNNFQNGOLEARKNDHFRFSSKIKPRLSHNVTLRKE 667

QY 860 TANELNS---HLOELNSPENTVAVELTDHKKDKKNCIIIVYVNPNTVATINLPS 911

Db 668 NFIERNDDITMLEDMMYNEESRFLAFMLHD---ONGDIYIATAFNAHHFSTIKTALPS 719

```

RESULT      6
US-09-463-238-14
: Sequence 14, Application US/09463238
GENERAL INFORMATION:
APPLICANT: Edwards, Elizabeth A
APPLICANT: Smith, Allison M
APPLICANT: Bustos Guillen, Regia
APPLICANT: Martin, Catherine R
APPLICANT: Plant Bioscience Limited
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 97.118
CURRENT APPLICATION NUMBER: US/09/463,238
CURRENT FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: PCT/GB98/02280
PRIOR FILING DATE: 1998-07-30
PRIOR APPLICATION NUMBER: GB 9716185.5
PRIOR FILING DATE: 1997-07-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 838
TYPE: PRT
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: SITE
LOCATION: (274, 276)
OTHER INFORMATION: Xaa = unknown amino acid
US-09-463-238-14

```

Query Match	10.0%	Score 504	DB 5	Length 838
Best Local Similarity	26.2%	Pred. No. 2.9e-28		
Matches 186	Conservative 97	Mismatches 262	Indels 164	Gaps 29

QY	338	YYSDDD--GNTYQKATTEFW---APTSPQVNVLLYDSATSGVTKY---PMTASGHC	369
Db	122	FYQKRPRLPFGATATNDGVNRFVFGNATTAATLCLITLSDRLPERKRYEQIILDPPLAKRTGD	101
QY	390	WEATVNQNLNNMYWYEVYGO-----GSTRTAVDEYATAIAPNGTRGMI---VDLA	438
Db	182	VMHFELKGFENMLYGRKFGDKFCEBGGHYFDSQIYLDYAAAIYRSRGYGVLPEDDC	241
QY	439	KTDVAG-----WNSKHTTPKNIEDEVYLENDYVDEFSIDPNSGKNKGTALYLT	487
Db	242	WPPNAGVPRCEFWISLWYEGDRLPHEVSHRELIXENHVGFTIHESSEKRYPGYLGVV	301
QY	488	EKGTGPGDNVATGIDSLKOLGITHVOLAPFASNSVE-----TPPTDNNGYDPRNY	540
Db	302	EK-----LDHLKELGVNCLIELMPCHEPHELEYYSYNSVLDGKRYNPGYSTVNF	350
QY	541	DVPEGQATNANGN---ARIKEREMYLSLHREHIGVNMDDVYNNHTFATQ---ISDFDK	593
Db	351	FSPMGRYSAGLSNCGAINEEFYULYKAHKRQIEIMDVYNNHTEGNENPILSFRG	410
QY	594	IVPEYRYRTMIOVILPTDOYLEMK-----LXAERPGVOKFIIDSLKRWVNEYHIDGRFD	648
Db	411	IDNSVFT-----TLAPGGEFYNTSGCGNTNRCNNPIYRQFV-MLRWYETEMHVGHRFD	464
QY	649	LMALLGKDTNKAASELHAIN-GIALYGEPMWGT-----SALPDDQLTLTKGA---	696
Db	465	LASIL-----TRSSSNWNAVYVYNSIDGQVITGTGELTSPILDMISNDPL-RGVKLI	518
QY	697	-----QKGM-----GV-AVFNDRMLRALDGNVFDSSAGCFATGATGLTIAIKNGVE	741
Db	519	AEAMDCGGLYOVGEPHMGWIMSENGNGYRDMV-----RQFIKGTDFSGAFACELC	569

```

0Y 742 G$IN$FT$----$P$G$T$IN$V$T$H$D$V$T$L$M$D$K$L$A$L$S$P$N$D$S$E$A$D$R$I$K$M$D$----- 786
Db 570 G$P$N$L$Y$G$G$R$K$P$M$S$IN$F$V$C$A$H$D$F$T$L$A$D$T$T$Y$N$K$H$L$A$N$G$E$D$N$K$O$E$N$H$N$S$N$C$E 629
0Y 787 -E$L$A$A$V$-----T$S$G$V$P$M$O$G$E$E$M$L$R$X$K$G$D$N$D$Y$N$A$G$A$V$N$E$D 829
Db 630 E$E$F$A$S$I$F$K$L$R$K$R$O$M$R$N$F$L$C$L$A$V$S$G$V$P$H$Y$M$G$E$D$G$H$T$K$G$G$N$N$Y$T$C$H$D$N$Y$N$T$R 689
0Y 830 W$R$K$O$Y$P$D$V$P$N$Y$G$L$I$-----H$L$R$H$P$A$P$R$M$T$A$N$F$N$H$L$O$F$L$N$S$P$E$-----N 876
Db 690 W$D$K$O$E$S$S$D$F$L$R$F$G$L$M$T$K$F$R$E$C$E$S$I$G$I$D$-----G$P$T$A$E$R$L$O$W$H$G$H$P$R$T$D$M$E$T$S$R 745
0Y 877 T$A$Y$E$L$T$H$V$N$K$D$K$G$N$I$Y$V$N$P$K$Y$A$T$I$N$L$P$S$G$-----K$A$I$N$T$S$K 922
Db 746 F$A$F$L$V$D$K$V$---G$E$L$I$A$F$N$A$S$H$L$P$V$T$T$L$P$D$R$P$G$P$V$D$T$K 790

RESULT 7
US-09-463-238-15
: Sequence 15, Application US/09463238
: GENERAL INFORMATION:
: APPLICANT: Edwards, Elizabeth A
: APPLICANT: Smith, Allison M
: APPLICANT: Bustos Guillen, Regia
: APPLICANT: Martin, Catherine R
: APPLICANT: Plant Bioscience Limited
: TITLE OF INVENTION: Starch Debranching Enzymes
: FILE REFERENCE: 97.118
: CURRENT APPLICATION NUMBER: US/09/463,238
: CURRENT FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: PCT/GB98/02260
: PRIOR FILING DATE: 1998-07-30
: PRIOR APPLICATION NUMBER: GB 9716185.5
: PRIOR FILING DATE: 1997-07-31
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 15
: LENGTH: 397
: TYPE: PRT
: ORGANISM: Solanum tuberosum
US-09-463-238-15

Query Match 5.8%, Score 293.5; DB 5; Length 397;
Best Local Similarity 25.5%; Pred. No. 4.8e-14;
Matches 96; Conservative 66; Mismatches 142; Indels 73; Gaps 15

0Y 623 P$M$O$K$F$I$D$S$L$K$Y$W$N$E$H$I$D$F$R$D$I$A$L$L$K$D$F$-----M$S$K$A$S$E$L$H$A$I$N$P$G$I$A 673
Db 36 P$Y$M$E$L$L$I$E$S$L$H$M$W$T$E$H$Y$D$G$F$R$D$L$A$S$V$L$C$R$G$D$G$P$P$I$N$A$P$P$Y$K$A$S$K$-D$S$Y$L$S$R$C$K 94
0Y 674 L$Y$G$E$F$W$T$G$---T$A$L$P$D$D$L$L$T$G$A$O$G$M$G$V$A$F$N$D$L$R$M$A$L$D$G$N$V$D$S$A$O$G$-F$A$T$G 728
Db 95 I$I$A$E$W$D$C$G$G$L$Y$V$K$F$P$M$D$R$M$A$E$M$K$R$-----Y$R$D$D$I$R$F$I$G$---D$A$G$K$G$F$A$T$R 145
0Y 729 A$T$G$L$D$A$I$K$N$G$E$G$S$I$N$D$T$S$P$G$E$T$I$N$Y$T$S$H$D$Y$T$L$M$D$K$L$A$L$S$N$P$-N$D$S$E$A$D$----- 781
Db 146 I$A$G$S$A$D$L$Y$R$-----V$N$--K$R$K$P$Y$H$S$V$N$V$I$A$H$D$E$F$L$Y$D$L$V$S$Y$N$N$K$H$D$A$N$G$E$G$D$O$C$ 197
0Y 782 -----R$I$M$D$E$L$A$Q$V$V$M$T$S$G$V$P$M$O$G$E$E$M$L$R$X$K$G$D$N$D$Y$N$ 819
Db 198 N$D$F$S$M$N$G$I$E$G$E$T$S$D$A$N$I$N$A$L$R$O$M$K$N$F$I$H$A$L$V$S$O$Q$T$P$M$L$G$D$E$D$G$H$T$R$Y$G$N$N$S$Y 257
0Y 820 N$A$G$D$A$V$N$E$F$D$M$R$-K$A$O$Y$P$D$V$P$N$Y$S$G$L$I$H$L$R$D$P$A$P$R$M$T$A$N$F$N$S$-----H$L$O$F$L$N$P 874
Db 258 G$H$D$T$A$I$N$F$O$M$Q$L$E$A$R$K$N$D$H$R$F$S$K$M$K$F$L$S$H$N$V$L$K$E$N$F$I$E$K$N$D$I$T$W$L$E$D$N$Y$N$E$E 317
0Y 875 E$N$T$A$Y$E$L$T$D$H$V$N$K$D$K$G$N$I$Y$V$N$P$N$K$T$V$A$T$I$N$L$P$S$G$K$M$A$I$N$A$T$S$G$K$E$S$T$L$G$A$E$G$S 934
Db 318 S$R$E$L$F$M$L$H$D$-----G$N$G$D$I$Y$A$F$A$N$H$P$S$I$K$T$A$L$P$S$P$-P$N$R$S$Y$R$V$V$D$T$N$L$K$S$P$D$F 371
0Y 935 V$O$Y$G$I$S$M$M$I$L$H$O$E$V$S$P 951
I : I : I : I :

```

```

DB 372 V-IEGSGISETYDVAP 387

RESULT 8
US-09-463-238-4
; Sequence 4, Application US/09463238
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Allison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GS98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-4

Query Match 5.3%; Score 270; DB 5; Length 878;
Best Local Similarity 22.4%; Pred. No. 6.9e-12;
Matches 148; Conservative 91; Mismatches 258; Indels 164; Gaps 29;

QY 256 INULTPAGCAHATFYSYIPSTHAUVYPTINPNMADLQVESGKRD-LVYTLGSDPDVSHTL 314
DB 172 VALDEASIALPRTYISFYKMSQLYSDMENS-----EIRSHRNTEYVPPGLSGHPAPLGI 226
QY 315 SIOTDGYAKQYIPRNVLSQYYSGGDLGNTYQKATTFKVMAPTSTOVNLLYDSAT 374
DB 227 SFQPGP-----SVNFALFSRARSAYVLCIYDOI- 254
QY 375 GSVTK-----IVPRTASGHVMEATVNONLENNYTFE-----VVGSGSTRANDPYATA 424
DB 255 -SVERPSLEIDLDPRYINRSGDIMHALDCLSPFKTYGYRCKATTSGKGL-VLDDPYAKV 312
QY 425 I-----APNGTR---GMIVDLAKTPDAGNSDKHITPK-NIDEVIYENDVADFSLDPS 475
DB 313 IRRVIRPGSGSIRIRKYLGELCELCGIDYMGSD--VPPSLPMKLLITRLANTYQFLKDKSS 370
QY 476 GKKN--KGKYLALTEKGTGPDNVKGTGIDSLKOLGITHVQLMPVFASSVDETDPTODNW 533
DB 371 KLPDDLACTGFFSGISEKW-----HHFKDLGVNAMLLEPIF-----PFDEQK 410
QY 534 G-YDPRNVYVPGGOYATNANGNARIKEPEKFWLSLHREHIGNMNVYVNHPE-----ATQI 588
DB 411 GYFFPHHFFSPGNNWGPSSDPLSAIKSMKDMYKLLHANGIEVLEVFTHHDAEDAPLANV 470
QY 589 SDPDRIVPEYRYRTMIOVLIPTDOYLEMKLAEKRPVMOYFIIDSLKYWVNEHIDGFRED 648
DB 471 DNFYSYCIKGGQYLN-----IYNALNCLNYPYVQGMILDCRLHMYTEHIDGFVAF- 518
QY 649 LMAALLGKDTMSKAASELAHINPGIALYGEPRWTGGISALPDDOLLKGAOKGKGVAVFNDN 708
DB 519 -----VNASSILKGFN-GEILSRPL--VEALAFDPLK-----VKMIADN 557
QY 709 ---LRNALDGNFED-----SSAGCFATGATGTIDAIKNGEGSINDPVS--S 750
DB 558 WMPLRNDSKEMFLPHMRMAELNMRFCDIDRFNG-EDLLSNLATRCLGSSDIDTAGRG 616
QY 751 PESTINVTYVSHONTYLMKIAL-SNPNDSEAD-----RIKMDLAQAV 792
DB 617 PAFSFRYIARNSGTLVLDVYSPSSNEVASLWNGOGEGATTNNIVLERRLKOVNPLFI 676

```

QY 793 VMTSGVPMOGGEEMLRXKKGNDNSYNAGDAVNEFDMS--RKAQYDPVENYISGLTHLR 850  
Db 677 LFTSLGAVPLNMGDECCOSSGSGPP---AYDARKSLGMNTLKTGCTQIAQFIISFLSNLR 732  
QY 851 L 851  
Db 733 M 733

RESULT 9  
US-09-463-238-19  
; Sequence 19, Application US/09463238  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Elizabeth A  
; APPLICANT: Smith, Alison M  
; APPLICANT: Bustos Gullien, Regla  
; APPLICANT: Martin, Catherine R  
; APPLICANT: Plant Bioscience Limited  
; TITLE OF INVENTION: Starch Debranching Enzymes  
; FILE REFERENCE: 97.118  
; CURRENT APPLICATION NUMBER: US/09/463,238  
; CURRENT FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: PCT/CB98/02280  
; PRIOR FILING DATE: 1998-07-30  
; PRIOR APPLICATION NUMBER: GB 9716185.5  
; PRIOR FILING DATE: 1997-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 883  
; TYPE: PRT  
; ORGANISM: Solanum tuberosum  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (603, 606)  
; OTHER INFORMATION: Xaa = unknown amino acid  
US-09-463-238-19

Query Match 5.3%; Score 269; DB 5; Length 883;  
Best Local Similarity 22.4%; Pred. No. 8.2e-12;  
Matches 148; Conservative 91; Mismatches 258; Indels 164; Gaps 29;

QY 256 INTLVPAGAAIVTSTYPSHTAVYDITINPNADQVESGKTD-LVYVTLGDEPDVSHTL 314  
Db 177 VEIDPEASLAPFYISFYMKSQLVSDMENS---EIRSHRTNPFVVGSLSSGHPAPLGI 231  
QY 315 SIOTDGTQAOQVIRVNLNSQYVYSGDDLGNVTOKATTEFKVMAPTSTOVNVLVYSAT 374  
Db 232 SFQPDG-----SYNFALESRSARSVVLCIYDI- 259  
QY 375 GSVTK-----IVMTASGHQWEATVNOLENNYVWE---VTGCGSTRTAVDPYATA 424  
Db 260 -SVKPSLEIDLDPIYINRSGDIMHAALDCSLPFKTYGRCATSTSGGEL-VLDDPYAKV 317  
QY 425 I-----APNCTR---GAVIVLAKTDPAGMNSDKHITPK-NIEDVYIEMDVRODSIDPNS 475  
Db 318 IRRVIRPQSGSEIRPKYVIGELCEPGYDWSGD--VPPSLPEKXLIIRLANVTOFTKDKSS 375  
QY 476 GKNK--RGKTLALTEKGTGKPDNKTGIDSLKQGLITHVOLMPVEFASNSVDETPDQDNW 533  
Db 376 KLPPDLAAGTSSGISEK-----IHFKDLGVNALLPEF-----PDEQK 415  
QY 534 G-YDPNNYDPEGOYATNANCNARIKEFEKENVLSLAREHIGVNDVYVNHTE---ATOI 588  
Db 416 GPYPMHFESFGNNYGPSGPDLSAIKSMKDMVKKLHANGIEVELEVFTHTAEADAPLMNV 475  
QY 589 SDFKAIPEYIYRTMIOVITPTDOVLDMKLAERPMYQKFIIDSLKATVWNEYHIDGRFD 648  
Db 476 DNFSYCKGGOYLN-----IONALNCNPIVQOMLITLRRHVIIEFHIDGVEF- 523  
QY 649 LMLLIGDTMSKASELHAIIPGIALYGEPMWTGSTALPPDQDLITKGAQGMGAVFVNDN 708

Db 524 -----VMSLLRGEN-GEILSRPPL---VEAIADPILSK-----VKMATDN 562  
QY 709 ---LRNALDGNVFD-----SSAOGFATGATGTLDAIKNGVESINDFTS--S 750  
Db 563 WNPPLTNSKENLFPHMRMAEINMRFCODLRDLRG-EGLLXNLXTPRLCGSGDIFACGRG 621  
QY 751 PGETINVTSHDNTLMDKIAL-SNPNDSEAD-----RIKDELAAQAV 792  
Db 622 PAFSFNYIARNSGILTLDVLSFSSNEVASSELNWCOCGATNNIVLERRLKQVRNFLFI 681  
QY 793 VMTSGVPMOGGEEMLRXKKGNDNSYNAGDAVNEFDMS--RKAQYDPVENYISGLTHLR 850  
Db 682 LFTSLGAVPLNMGDECCOSSGSGPP---AYDARKSLGMNTLKTGCTQIAQFIISFLSNLR 737  
QY 851 L 851  
Db 738 M 738

RESULT 10  
US-09-739-449-8397  
; Sequence 8397, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 8397  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-8397

Query Match 5.2%; Score 265; DB 5; Length 379;  
Best Local Similarity 23.1%; Pred. No. 4.4e-12;  
Matches 97; Conservative 71; Mismatches 148; Indels 104; Gaps 17;

QY 504 LKQGLITHVOLMPVEFASNSVDETD---PTODNKGYPNRYDPEGOYATNANGNARIKE 559  
Db 21 LKKIGYDAVEIMPIYA--WIDERHLPLIGLSNMGVNPVGFIALDPRLCPGCG---VRE 73  
QY 560 FKENVLSLAREHIGVNDVYVNHTEFATQIS-----DFPKIPEYIYRTMIOVITPTDQ 612  
Db 74 LRDTVAALHMEGIVIIDLVFNHTGESDIEGSIISLGLDLTRFRHPQGRGVLY-NDT 132  
QY 613 VLENKLAERPMYQKFIIDSLKYVWNEYHIDGFRFDLMALIGK-----DTMSKASELHA 667  
Db 133 GTGVTACDHPYIQLVYDLSRHFVNLNANGYDGFEDLAPVLRGAAGFDMASELAAHMS 192  
QY 668 INPCI---ALYGEWGTGTSALPPDQDLITKGAQGMGAVF--NDMLRNALDGNVFDSS 721  
Db 193 -DPVLVRLRLIAEPWDIGPG-----GYQGNPESEFLEWMDRARDM----- 233  
QY 722 AOGFATGATGTLDAIKNGVESINDFTS---PGETINVTSHDNTLMDKIALSNPNS 778  
Db 234 -RRWRGDAGTGTGALANALSSSPISFNRHGLKRSYVFLAHDGFTLFDLVSHERRHNE 292  
QY 779 EADRIKDELAAQAVVMTSGVPMOGGEEMLRXKKGNDN-SYNAGDAVNEFDMSRKAQY 837  
Db 293 -----KNGEN--NRGHNDNSWNG----- 311  
QY 838 DVFNYYISGLHLRLDHPAFRTTANETINSLOFINSEPNVAVELTDHVKNKKGNTIIV 897  
Db 312 -----FEGL-----TDPAIYAARLADYKALLSTLFVSRGTLMLJTAGDEGSHSGOGNMPI 362

RESULT 11





QY 123 QIR-----NEKDAEDA-----AKPAVNA-----YLDASNOVLVKL 154  
DB 567 LASTLINEVNTADALOSFTVYKEDDDADANAIVAKDPTNAGAVSLIKKKGKGLVAT 626  
QY 155 SQTTLLEG---XSGFTVHDDTANKD-----IPVSVYDASLGOVTA 194  
DB 627 KKGDTVFFGLSODSGLTIGSTLNBDGLYKDTNEQIQVANGIKFTNAGSNGTIGIAN 686  
QY 195 VLAQTFOHI-FGSGDMAPDNHSTLLKVTNNLYQFSGDLPEGNYQYKVALNDSWNSPS 253  
DB 687 TARTTRDKIGFAGSDGAVD-----TNKPYLDQDKLQGVNKI-----723  
QY 254 DNINLTVPAGANHTFSTIATVYDTINNPNADLOVESGVKTDVTVLG---EDPDV 310  
DB 724 --TNTGIMAGKAIT--GLSPTLPSIDQSSR-----NIEGNTIQDKD 764  
QY 311 SHTLSID---TDGYAKOVIPIRNVLNSSOYVYSCDDLNTYQKATTFKWAFTSTOVN 367  
DB 765 SNASIDILINTGNLKN-----NNNPIDFVSTYDIDVANGNATTAITYTHDTANKTSK 818  
QY 368 LLYD-----SATGS-----VTKIVPMTASGHGWEATVNCNLE-----400  
DB 819 VVYDVAVNDFTTHLTCGHDNKKLGKVTTKLNTKSANGNFATNPNVNSDEBDALVNAKDIA 878  
QY 401 ---NMYYMEVTOGGSTRTAV-----DPYATAPNCTRGHIVDLATDP 442  
DB 879 ENLTLAKETHTTKGTADTLQFTYVKKVDENNADANAIYQOKNANNQVNTLTJLKE 938  
QY 443 AGWN--SDKHITPKNIDEVLEYEMDVDFSIDPNSCKM-----NKSKYLALTEKTKGP 494  
DB 939 NGCLNTKIDKNGTV-----TFGINTTSGLKAKSKSTLNDG---GLSKNTTGS 981  
QY 495 DNKATGIDSLKOLGITHVOLMPVAFASNSVDEDTQDNMGYDPRNVDYPRCOATNANG- 553  
DB 982 EOIOVGADGVK---FAKYNNGVYVACIDGTTTRITRDEIFGTGNSLDSKPHLSKDCI 1038  
QY 554 NARKKEKENVLSHREHIGVNNVDVYNNHFAQISDFDKIVPEYRTIMQVIPIPDV 613  
DB 1039 NAGCK---ITNIOGSELIQON---SHDAVTCGKIYD-----LKE-- 1072  
QY 614 LEMKLAEKRPVOKFIIDSLKYVNEYHI---DQFRF-----DLMALLGK 655  
DB 1073 LENKISSTAKTAQNSL-----HEFSVADEGNNFTVSNPYSYDTSKTSVITPAGE 1124  
QY 656 DYNKSAKSELHAINPGIALYGEPTGTSALPDQLLTKGAQKMGVAVFENDLRNALDG 715  
DB 1125 NGITTT---KVNKGVAVGIDQFKGLT---PKLTVCNNNGKGIIVDSQONQNTITG 1174  
QY 716 ---NVFD-----SSAG-----FATGATGLTDAIKNG---VEGSINDFTSSPGETI 755  
DB 1175 LSNTLAVNTDNKGSVRTTEGQNIKKDEKTRAASIVDLVLSAGFNLQGN-----GEAV 1226  
QY 756 NYVTSHDNNTLMDKIALSNPNDESEADRIKMDLQAQAVVMTSOCVPFQGGEMLRXKG- 813  
DB 1227 DFVSTYTVFAD-----GNATTAKVYIDDPISK---TSKYVYDVAVNDTTIEVADK 1275  
QY 814 ---GNDNSYNAAGAVNEPDSRKAQYDPVFNYSGLIHLR-LDHPRAFMTTANEINSHLO 869  
DB 1276 LGVKTTLTSTGTCANKFALSNCATGQALVKASDIVAHLMTLSGDIDJOTANGASQANSAG 1335  
QY 870 FLNRPENTVAYELTDH-----VNDKMGNIIVYNNPKVTATINLPS-- 911  
DB 1336 YVADGKAKVIYDSTDNKYQAOKNDGTVDKTEVAKDKL--VAQAQTGDTLQOMNVSVI 1393  
QY 912 GKMAINATSGKVG 924  
DB 1394 NKEGVNANKKOG 1406

RESULT 13  
US-09-739-449-12578  
; Sequence 12578, Application us/09739449  
; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 12578  
; LENGTH: 530  
; TYPE: PRF  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-12578  
  
Query Match 3.3%; Score 165; DB 5; Length 530;  
Best Local Similarity 20.3%; Pred. No. 7.1e-05;  
Matches 118; Conservative 69; Mismatches 172; Indels 222; Gaps 32;  
  
QY 415 RTAVDPYATAIAPNGTRG--MIV--DLAKTDPAGWNSDKHITPKNIEDVYEMDVDFPS 470  
DB 5 RTVRADPASNROEGRGPSPSLVNHDFAMKNP--NMKG-----RPMHEAVYVELHIGTFT 57  
QY 471 IDPNSGKKNKGKIALTEKGTGKPDNKTGIDSLKOLGITHVOLMPY--FASNSVDETD 528  
DB 58 -----PGCTFAAAEK-----LEVLDVGITTIIDMLPATFAGSR-----92  
QY 529 TDQNMGYD-----PRNVYPRCOATNANGNARIKEFEKENVLSLREHIGVMDVYVN 581  
DB 93 ---GNGIDGYLQSPQNDYGTDP-----ELKAFIDQAHGHIMVLDVYVN 135  
QY 582 HTEATQISDFKIVPEYVYRTIMQVIPIPDVLEMKLASRPVOKFIIDSLKYVNEYH 641  
DB 136 H-FGPACNTLQAVAPAFKKHE---TPMGAPADF---NRAEYRSPFLQNAFYWLETYR 186  
QY 642 IDGFRPL-----MALGKDTM-----SKAASELHAINPGIA 673  
DB 187 FDGLRIDADHLAGGDEVDLEMAREVKRTINRRVHVLIEDARNAAASPMTPMADGAI 246  
QY 674 LYGPWTGTSALPDD-----QLLTGKQKMGVAVFEND---NLRNAJQONVDFSS 721  
DB 247 LVDQWN-----DDPHNHYHVAITNEEGS---IYEDASRPYELMLRSL-----287  
QY 722 AQGFATGATGLDAIKNGVGSINDFTSS-----PGETINVTSHDNTLMDKIALS 773  
DB 288 ---ATGF---VYCGEPRPSRNFAASGSPSCHLPPHRFVNFPLHNH-----325  
QY 774 NPNDSEADRIKMDL-----AQAVYMTSOCVPFQGGEMLRXKGNDNSTNAG 822  
DB 326 ---DOAGNRLEGERLRALIPPLFGTLEAILLCPOTPLVFMGDE---HGSANP---373  
QY 823 DAVNEFMSRKAQYDPVFNYSGLIHLRDLHPAFRMTTANEINSHLOFLNSPENTVAYEL 882  
DB 374 -----FFFFFS-----DHPDH--NREDEIRRLKQAESFQG---EL 403  
QY 883 TDHVNKDKMGNIIVYNNPKVTATINLPSCKWA--INATSK 922  
DB 404 PPDASQ-----VMDPNDQ--HTMQSLTKLWTAETTEGR 436

RESULT 14  
US-09-739-449-9531  
; Sequence 9531, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23

```

; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9531
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9531

```

```

Query Match          3.2%; Score 161.5; DB 5; Length 532;
Best Local Similarity 18.8%; Pred. No. 0.00013;
Matches 115; Conservative 75; Mismatches 199; Indels 225; Gaps 28;

```

```

QY 460 VYEMDVDFSDIDPNSGKMKNGKATLATEKGTGKPDNVTGTGIDSLKLGITHVQLMPVFA 519
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 VYQVYPRSFQDTTGDGY---GDLACVTKR-----LPIYASLVDAIWLSPFT 45

QY 520 SNSVDETPDQDNMGVDRNRYDVEGOYATNANGNARIKEFEEMVLSLREHIGVMDV 579
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 46 -----TSPMKD-FGYDVSNVVDPMFCT-----LADFGLLAEARLGIWIDLV 91

QY 580 YNHT-----PATQISDEDKIVPEY 599
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 92 LSHTSDDHPWFVESRASNNKSDMYVWSEAKPDGTPNNMWSITGSGQMNDPTRMQY 151

QY 600 YRTMIQVITPTDQVLEMKLAEPRVOKFIIDSLKYWNEHYHIDGFRDLMALLKIDTWS 659
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 152 MHNFL-----TSQPDILNHNPEVEOELLNITRFWLKR-GVDGFRDITINFFHDL-- 200

QY 660 KAASELHAINPGIALYGPWVGTSALP-----DDQLITKGAKKMGVAAPFNOLRNA 712
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 201 -----ELRD-NPRLA---PERRNASTAPAVNPFNFQEHLYDKNRPENT--AFLRFRAY 248

QY 713 LD-----GNVDFSSAQGFATGATGLTDAIKNGVEGSINDFTS----- 749
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 249 LDEPDIANAAGEVGS-----QRLGE-IVGYTSGDDKQMCYAFELIA 291

QY 750 -----SPGETIYV-----TSHDNTLMKIALSNPN-----DSEADRIKMDLQAO 791
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 292 PDALPQVAVDQADFAARAPEGWACW--AFSNHDVVVRHVSRMCEHVEDKDAFAKLISA 348

QY 792 VVMTSGVPRFMOGGEBMLRXKG-----GNNS-----YNAQDAVN 826
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 349 LMTORGSVCYIEGELGLTEADIAFEDLDQPYGIQFMEFEFGGRDCRTPMWMDGHQA 408

QY 827 EFDWSRKAQY-----DYFNYSGLIHLRLDHPAFRMTTANEINSHL 868
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 409 GFSTSDKIMLRIPAKHQRAVASAOGCNEASVLEHYRRFLAFKQHPAF-----AKGI 461

QY 869 QFLNSPENTVAYELTDHVNKDKGN--IIVVYPNKTYATTINLPSGKMAINATSGKVGES 926
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 462 EFQPLEGCVLSY-----TRKIGNFTILCLFNLSATPANATLPGKNMEV-----LEGHG 509

QY 927 TLGOAEG-SVOYP 938
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 510 PASGLEGTVELP 522

```

```

RESULT 15
US-09-739-449-9662
; Sequence 9662, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514, 000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9662
; LENGTH: 533
; TYPE: PRT

```

```

; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9662

```

```

Query Match          3.0%; Score 150; DB 5; Length 533;
Best Local Similarity 18.2%; Pred. No. 0.0008;
Matches 97; Conservative 76; Mismatches 175; Indels 186; Gaps 22;

```

```

QY 460 VYEMDVDFSDIDPNSGKMKNGKATLATEKGTGKPDNVTGTGIDSLKLGITHVQLMPVFA 519
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 VYQVYPRSFQDTTGDGY---GDLACVTKR-----LPIYASLVDAIWLSPFT 46

QY 520 SNSVDETPDQDNMGVDRNRYDVEGOYATNANGNARIKEFEEMVLSLREHIGVMDV 579
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 47 SPMA-----MGYDVSDYCNVDPMFCT-----LADFLALMEARLGIWIDLV 91

QY 580 YNHT-----PATQISDEDKIVPEY 599
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 92 ISHTSDRHPWFVDSRASRINSKADWYVMANPRPDGTAPTNMLSVEFGPAMEMDGVYKQY 151

QY 600 YRTMIQVITPTDQVLEMKLAEPRVOKFIIDSLKYWNEHYHIDGFRDLM----- 650
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 152 MHSFL-----ASQPDLPFNHPEYQDAVLETVRFLWR-GVDGFRDITVNHYPFDKLL 202

QY 651 ----ALLGKDTMSKAASELHAINPGIALYGPWVGTSALPDQLLITKGAKMGV---- 702
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 203 RNPPLVDDSEFGLDASD--VNP-----YG-----MDHLVDKTRPENAVFLKRF 245

QY 703 -AVFNQNLNRLDGNVFD-----SSAOGFATGATGLT-----DAIKNGVEGSIND 746
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 246 RALDDYEGRAIVGEYGDGARSILKTVAAATSGNDKLMCYTFDLAGPDRSAKLIRGSVER 305

QY 747 FTSSPGETINVTYSHDNTLM-----DKTA-LSNPNSEADRIKMDLQOAVVMTSGVPR 800
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 306 F-----GKYVT-----DGWYCAAFSNHDVYRHLRSRSEGEBOQIRVAKLAIYVLSARGSI 356

QY 801 FMOGGEBM-----LRXKKG-----NDNSYNAGDAVNEFDM 830
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 357 CLYQGEELGLTEAELAFEDLRDQYGIHFPAFGRDCRTPMWETGKPNAGFSTAEKPM 416

QY 831 -----SRKAQYPD-VFNYSGLIHLRLDHPAFRMTTANEINSHLOPL 871
    |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 417 LPVPYVHAMQAVDAQDERKPDVLYNHYRAVLSFKRTGALRLDGDMPHRIKTNLDVL 470

```

Search completed: April 22, 2001, 09:09:21  
Job time: 697 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      April 22, 2001, 07:55:59 ; Search time 62.87 Seconds
              (without alignments)
              1045.000 Million cell updates/sec
```

Title: US-09-262-126C-2  
Perfect score: 5060  
Sequence: 1 MAKRLIYVCLSVCLVTWAF.....VPGISMILHQEVSPDHGKK 956

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Searched:      198801 seqs, 68722935 residues
Total number of hits satisfying chosen parameters: 198801
```

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1570	31.0	843	2	H72204	pullulanase - The
2	1084.5	21.4	717	2	H84057	pullulanase amyl
3	1039.5	20.5	718	2	G69585	pullulanase amyl
4	819	16.2	1072	2	A84112	alkaline amylopul
5	815.5	16.1	910	2	D75524	alpha-dextran end
6	797	15.8	962	2	T01688	starch debranching
7	793	15.7	986	2	T03760	alpha-dextrin end
8	787.5	15.6	1090	2	S11823	alpha-dextrin end
9	786.5	15.5	904	2	T04377	probable alpha-dex
10	771	15.2	964	2	S51324	pullulanase - spli
11	759.5	15.0	1096	2	A26879	alpha-dextrin end
12	721.5	14.3	1038	2	S38801	alpha-dextrin end
13	533.5	10.5	666	2	E71565	probable hydroclas
14	517.5	10.2	666	2	G81717	glycosyl hydrolase
15	505.5	10.0	707	2	S77094	glycogen operon p
16	497.5	9.8	718	2	S73088	glycogen operon p
17	487	9.6	776	2	S13470	isomylase (EC 3.2
18	487	9.6	776	2	A37035	isomylase (EC 3.2
19	480.5	9.5	789	2	T01331	probable isomylase
20	472.5	9.3	664	2	B72084	glycosyl hydrolase
21	471.5	9.3	720	2	B75540	glycogen operon p
22	465.5	9.2	818	2	T02221	probable isomylase
23	452.5	8.9	746	2	S74564	glycogen operon p
24	440.5	8.7	721	2	A70764	probable g1gX prot
25	412	8.1	659	2	A64119	glyx protein (EC 3
26	385.5	7.6	656	2	D82388	glycogen operon p
27	379.5	7.5	552	2	S19134	probable alpha-gly
28	363	7.2	657	2	BVECGX	glycogen operon p
29	360	7.1	716	2	C83375	probable glycosyl

30	343.5	6.8	1475	2	A44765	alpha-amylose (EC
31	333.5	6.6	1481	2	S28669	amylolipulanase p
32	320.5	6.3	600	2	H75516	probable maltotriolig
33	292	5.8	1203	2	S27545	pullulanase - Therse
34	284	5.6	558	2	JC5135	pullulanase (EC
35	278	5.5	551	1	S48130	alpha-amylose (EC
36	274	5.4	2032	2	I38917	cyclomaltoextrinase
37	272	5.4	561	2	S73087	hypothetical prote
38	255.5	5.1	720	2	E72074	alpha-amylose (EC
39	254	5.0	630	2	D70363	1,4-alpha-glucan b
40	251.5	5.0	574	1	A42950	cyclomaltoextrinase
41	246	4.9	556	2	JC5132	alpha-amylose (EC
42	246.5	4.9	583	1	J50673	neopolipulanase (EC
43	245.5	4.9	705	2	D75345	neopolipulanase (EC
44	244.5	4.8	788	2	S70079	probable 1,4-alpha b
45	243	4.8	566	1	S34731	amylose (EC 3.2.1.1

## ALIGNMENTS

RESULT 1  
H72204  
pullulanase - Thermotoga maritima (strain MSB8)  
C:\GenBank\Blast\Blast1

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C;Accession: H72204

R. Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwynn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, J.R.; Miller, M.; Steuart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A;Reference number: A72200; MUID:99287316

A;Accession: H72204

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-843 <ARN>

A;Cross-references: GB:AE001821; GB:AE000512; NID:g4982421; PIDN:AAD36907.1; PID:g498

A; experimental source: strain MSB8

C;Genetics:

```

Query Match Similarity      31.0%; Score 1570; DB 2      Length 843;
Best Local Similarity      36.6%; Pred. No. 6,9e-82;
Matches 349; Conservative 145; Mismatches 291; Indels 168; Gaps 22.

QY      34 TTTIHYECPADYDYPMSLAMPKRD---GGGAEDYFNPDPADSEFGAVASADIPGNPSOVGI 90
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      21 TTTIVHHIRYDCKYKGMNIMINPVPYVSEGAAYQFT-GEDEFGKAAVAVKLLPMDLTKVI 79
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      91 IVRTQDW-TRDVSADRYIDLKSGN-EVWLVEGNSQIFVNEKDAE-----DAKPAVSNA 142
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      80 IVRLNMQAKDYAKDRFEIKDGRKAEVWILQVEEIEIFYEKKPTSPRIFFAQARSKNVIEA 139
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      143 YL----DASNOVLVLSQPLTIGEGKSGETTYHDDTANKDIPVTSYKSDASLGODVTAIVLAG 198
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      140 FLTNPVDTKKKKLEFYT-----VDGKEIIPVSRVEKA-----DPT----- 173
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      199 TFGHIFGSGDMAPODHHSTLLKKVTNNLQFSGDLEPGNGQYKVALNDSNNNSYPSDNIWL 258
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      174 -----DIDVTNT-VRIVISESL----- 189
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      259 TVPAGAHVTFSYIPSTHAVYDTINNPNADLOVESGVKTLVTLGEEDPVSHTLSIQT 318
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      190 -----KEDDRKDYELL 202
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      319 DGYQAKQVILPRNVLNSOYYYSGGDLGNTYQKATTFKYMABTSTQVNTLVLYDSATGSVT 378
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||
Db      203 EGYKAPAVLWMEILD--DYTYDG-ELGAVAYSPKRTIFRWKSPVSKVVKYLLFRKNGEDTLP 259
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

QY      379 -KIVMTWISGCHVEMATYQNLNNNTYKTEVTVGGSGTRFAVDPAVYATAIPNGTGRGIYDL 437
      |||:||||| : |||:||||| : |||:||||| : |||:||||| : |||:|||||

```



Db 54 EAKKY-VCVSD-----HP-----VTFGI---HCY-----RASS 79  
 QY 294 GVKTDLVTLTGEDPDVSHLSIQDGYQAKOVIIPRNVLNSQYVSGDDLNTYQKAT 353  
 Db 80 GHTKDL-----QIGAVI--RTAAPDEFYDQ-ELGAVYTAQHT 115  
 QY 354 TFKYVAPSTOVNVLVYSATGSGVTKIYVPTASGHWAEATVNONLEWYMYEYTGGS 413  
 Db 116 VFYVAPAASTAAVYK--SHPNKSGRTFQMTREKGVYAVVTGDLHGYEYLFICINNSE 173  
 QY 414 TRTADPAATAPARGTGEMIVDLAKTDPAGWNSDKHTP--KNIDEVIEYEDVDFSD 472  
 Db 174 WMEYVDQAKAVYNGEGVVL---RPQMKTP--IKPSSHVDAVYETHELDSETH 228  
 QY 473 PNSGKMKGYALATLTKGKTPGPDVKTGIDSLKQLGITHVQIMLPVAFNSVDEPTPD- 531  
 Db 229 ENSGINKGKTLALTEPTDYOANGSSGLAYVKELGVTHVELLPVNDFAQVDEKPLAY 288  
 QY 532 MWGIDPRYDVEGOYATNANG--NARIKEFKEMVLSLREHIGVMDVYVNTFATQISD 590  
 Db 289 MWGNPILHFFAPESGYASNPDPQTRKTELKOMINTLHOGLRVILIDVYFNHVYKRENSP 348  
 QY 591 FDKIYPERYPRMIOVIIPDQV--LEMKLXAEPRVOKFIIDSLKYVWNEYHIDGFRDL 649  
 Db 349 FEKTYPGTFER-HDECGMPSNGTGVGNDIASERRMARKFIADCVYMWLEETVNDGFRDL 407  
 QY 650 MALLGKDTMSKASELHAINPGIALYGPMTGTSALPDQDLITFGAOKGMGAVFENDL 709  
 Db 408 LGILIDITVLWKEKATKAKPGILLFGEGWDLATPLRHEOKAALNAPRMPIGPFNDMF 467  
 QY 710 RNALDGNVDSAGCFATGATGLDIAIKNGVEGS-----INDETSPGETTINYVTSNHY 764  
 Db 468 RDVAGNPFHLKATGFALNGESAQAQVMHGISGSKRALAPVPEPSQINYYESHDNH 527  
 QY 765 TLMQKIALSNPNDSPADIKMDELQAAYVMTSGVPRFMOGSEMLRXXGNGNNSYNAGA 824  
 Db 528 TFMDSALSLPOENDSRKRSRORLAVATILLAQVPTFHSGGEPRTTQOGVNSYQSSDS 587  
 QY 825 VNEPDSRKAQYPDVFNYSGLIHLRLDPAFRMTTANESHLOPLNSPENTVAYELTD 884  
 Db 588 INQDMDRREFPKEDVYIRLRLSLKKAHPARLRSADIOHLELCITLKEHLIAYRLAD 647  
 QY 885 HVNDRKGNIIYVYVNPKTATVATINLPSG--KMAINATSGKGESESTLGAQESGVYVPGISM 942  
 Db 648 LDEVDKMDIIVIHSHASPSVEMRLPNDIPYRLLCDPSGFORDPF--ELIKTVAVANGIGT 705  
 QY 943 MILH 946  
 Db 706 VILY 709  
 RESULT 4  
 A:84112  
 alkaline amylopolyluanase BH3697 [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence revision 01-Dec-2000 #text change 31-Dec-2000  
 C:Accession: A84112  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20263314  
 A:Accession: A84112  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1072 <STO>  
 A:Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BA07416.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BH3697

Query Match 16.2%; Score 819; DB 2; Length 1072;  
 Best Local Similarity 23.6%; Fred. No. 7.5e-39;

Matches 296; Conservative 165; Mismatches 402; Indels 294; Gaps 51;  
 QY 2 AKKLIYVCISVLVLTMAFNVNG--OSAHADGNTTIIYVFPACDQOPWLSMM----- 55  
 Db 6 ASKSVLFRVILVLLSLFPIRPMHEANADTEADLRIRHYQSSDOQFENGLMLNGVRY 65  
 QY 56 PRD---GGAEYDFNOPADSFQAVASADIPGNSQV--IIVTQDWTKDVADRYIDL-- 109  
 Db 66 PSENNPSCGTPPEPEKMTD-YGAVYDIELKENAENIGFLVMTATGDKD-GDKKHVQLLS 123  
 QY 110 SKGNVWLVEGSGQF-----YNEKDA-----EDAAPAAS----- 140  
 Db 124 PEMKEIWLKESGDVFLERPVDLPHTYRIHYERSDOSYHNGKLTWKVAPSEHEMPSG 183  
 QY 141 -----NAVILDA-----SNQVLKLSQPLTLAGXSGF--TVHDTANKDIPVTS 182  
 Db 184 AMPFDEQTSRGATILVDLADANEI-----GLVLVNENKDKD----- 223  
 QY 183 VKDASLGQDVT-AVLATGFQHI FGSDMAPDNHSTLLK-----VTNNLYQFSGDLPEG 235  
 Db 224 -----GDKTPADLA-----NHPOLFIRGQDGYVYTNPYQTIEGLVSG 262  
 QY 236 NYQYVALNDS-----W--NNSYPSDINILTVPAGCAHVFSTYIPSHAYVDTINNN 286  
 Db 263 EL---ISLTEMELRFTSTIWLSEDDLAEQILTL-DTNGASVAFDHV-----TTH--- 307  
 QY 287 ADLQVESGVKTDLVTLTGEDPDVSHLSIQDGYQAKOVIIPRN--VLNSQYVSGDDL 344  
 Db 308 -----DEKTYILHGSFDAL--LAPYHTPADKTVVARACHRLIDALYGTDEIG 354  
 QY 345 GNTYQKATTFKVMAPTSTQVNVLLY--DSATGSVTKIYVPTASGHWAEATV----- 396  
 Db 355 AVLHEDGSSTLFWPSPLADHVSVLVYDKDOYRIIKDDIDMTKGGQVYQVLTLSHNTGL 414  
 QY 397 ONLEWYVYEVYTGCGSTFAVDPAATAPRAGTR--MIVDLAKTDPAGWNSDKHTP- 453  
 Db 415 DNLAGYVYHYKLERGDTGLDIPYAKASMAANNESYSIGKAAVYDPSIGPELDPAFI 474  
 QY 454 ---KNIDEVIEYEMVDRDPSIDPNSGMKNK-----GKYALATREKGTGPDVKTGIDSLK 505  
 Db 475 EGFEREDALILEVYVHRTDTP--GIEBELTAQCTFAPSVDK-----LDYIE 521  
 QY 506 QUGITHVQIMLPVAFNSVDE-----TDPTQDNNGYDPRNYDVEGOYATN-ANG 553  
 Db 522 DLGVTHIQLLPMSYXWYWGDELANHERELAYSSGNNYNGCYPHSYFSLGYSSENPED 581  
 QY 554 NARIKEFKEMVLSLREHIGVMDVYVNTFATQISDPEKTYPERYPRMIOVIIPDQV 613  
 Db 582 ELRIKEFKHLIDEIHRGKGVILDVYVNTFAVEI--FEDLAPNYYH--FMADAGTPRTSF 638  
 QY 614 LEMKLAERPVOKFIIDSLKYVWNEYHIDGFRDMLALLGKDTMSKASELHAINPGIA 673  
 Db 639 GCGRLCTTHKMSRRLVDSITYVVEEYKVDGFRFDMGCHDAETIOIALDRAKELANPILV 698  
 QY 674 LYGEPR-----TGTSALPDQDLITFGAOKGMGAVFENDNIRNALDGNVDSAGCFAT 727  
 Db 699 MICEGRRTYVGBEHGCDIMPADQ---DMWQHTESSVGVFDEFERNEKLSGFSGCGRPTIT 755  
 QY 728 GATGLTDAIKNGVEGINFT--SSPGETTINYVTSNHYTLMKIALS---NRPDSADADI 783  
 Db 756 GGARNIEQIFDNLTAOPNHNFTATNPGDVVPYIEAHNDLTLHDVIAOSIRKDPYHOEELH 815  
 QY 784 KMDLQAQAVVMTSGVPRFMOGSEMLRXX-----CGN----- 815  
 Db 816 KRIRLGNMVLTLAAGTAPFLHACQEFORTQOFADTGOAPRYKSTYTDENGRFVYPYI 875  
 QY 816 DNSYNAGDAVNEFDMSR---KAQYD--VFNYSGLIHLRLDPAFRMTTANESHLO 869  
 Db 876 HDYSDSDAIKRFDMCKATNKERYPIHNLTRTYTACLIELRSTNAFTGTLVDQAVNR 935  
 QY 870 FLNSPE-----NTVAY-----ELTDH-----VNRKDKGN 893  
 Db 936 LDIPEISFTDLVIGYRANSTDNLTGYVYVNVADTKERTLSTLDLTGEGVIVDSDEAG- 994

OY 894 IIVYNN-----AKTATINLPSCMAIMATSGVSGESTLQAGSVOVPCIS 941  
 Db 995 IVAVSDPSGDLTPNETTIDPLTAVIRVGGDFQENPDHKGKG-GKAKGRPDHAGOP 1053  
 OY 942 MMLH--QEVSPDHGKK 956  
 Db 1054 GPPHAGEKCKPDMASK 1070

RESULT 5  
 D75524  
 alpha-dextran endo-1,6-alpha-glucosidase - Delnoccocus radiodurans (strain R1)

C:Species: Delnoccocus radiodurans  
 C:Date: 03-Dec-1999 #sequence, revision 03-Dec-1999 #text, change 17-Mar-2000  
 C:Accession: D75524  
 R:Miller, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Mc  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Delnoccocus radiodurans R1.  
 A:Reference number: A75250; MID:20036896

A:Accession: D75524

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-910 <WHI>

A:Cross-references: GB:AE001900; GB:AE000513; NID:96458079; PIDN:AAF09983.1; PID:9645808

A:Experimental source: strain R1

A:Genetics:

A:Gene: DR0405

A:Map position: 1

Query Match 16.18: Score 815.5; DB 2; Length 910;  
 Best Local Similarity 26.9%; Pred. No. 9.2e-39;  
 Matches 260; Conservative 126; Mismatches 332; Indels 247; Gaps 34;

OY 87 OVGIIIVTQMTKDVADRYIDLSKGNEMVLVEGNSQIFNEKDAEDAKPAVSNAFLDA 146  
 Db 7 KLGIVIVKGP-DKDPGPDGLITREQGMQMLVSGNPSTTRPSTR-----VRIVGDL 58  
 OY 147 SNOVLVLSQPL-----TLEGXSGFTVHDDTANKDIPVTSVKDASLQDVTAVLACTFQ 201  
 Db 59 KQOQAIMLSRLIVVKPALAQPCAMLTILH-----VSTATATLKGAD----- 99  
 OY 202 HIRGSGMAADNISTLLKKTNNLYOFSGDLPEGNTQYKVALDMSNNSTPSPNINILTP 261  
 Db 100 GVSQGDYLA-----LEVEEGCL-----TP 118  
 OY 262 ACGAHVTFYIPSTHAYVDTINPNADLOVESGVKTDLYVTIGEDPDVSHTLS-----IQ 317  
 Db 119 ALQAKAPYL-----AGYKL-----LQIRADRAKLPTALSGQ-LAVSSVLEPGKVLD 164  
 OY 318 TDGYQAKOVIIPKRVNLSSQYVSGDDLGNITTYOKATTFKVPASTPOVNVLLYDSATGSV 377  
 Db 165 ATGQVQFAMAL-----DDLYAYAG-PLGYVMQGNKPTVRLMAPTAQDVK-LHVSRRPDGTQ 216  
 OY 378 TKIVPMKASHGVWEALVNONLEMYMYTEV-----TGCGSTRTAVDPTATAIAPNGTR 431  
 Db 217 ERTLPMTKRTQGVWATASGDSMKGLAYRFVRFAPSTGNIETNLMTDPSVALTRNSTH 276  
 OY 432 GMTVDLAKTD--PAGMNSDHIITPKNIEDVIEYEMDVRODS-IDPNSGMKNKKCYALATE 488  
 Db 277 SVFRDLSDNSQKAPGNALKKPALRSFADLSFTEHLILRDSAADASVPAERGTLLAFTR 336  
 OY 489 KGTGKPDNVKGTIDSLKQLGITHVOLMPV----- 518  
 Db 337 PGSDG-----MTHIRALAGACILKAVHLLPTFDIATINEDKQWQTPQDLQFGPNSERQK 392  
 OY 519 ASNSVDETDPTQDNMGDPPKNYDPEQYATNANGNARIEFEKEMVLSLREHIGVNMV 578  
 Db 393 AVAAVADADAY--NMGYDPYHSHVPEGSYAUNP--AERTKEYROMVNSLMAAGILRVQDV 448

OY 579 VYNTFTAT--QISDFDKIVEYVYRTMIOVIPTDQVLEMKLAERPMQKFIIDSLKY 635  
 Db 449 VFNTHTAASGAQSVSLDKIIVPGYYHRLNNGVSNSTCCS-NTATEHTMRLAVDTLVL 507  
 OY 636 WYNEYHIDGFFRFDLMLLKGDTMSKASELHAINPG-----IALYGEFTGTSTALP 687  
 Db 508 MAREKYKDFRFDLMLLKGDTMSKASELHAINPG-----IALYGEFTGTSTALP 687  
 OY 688 DDQLLTKGAOK-----GMGAVFENDLRLNAL-DGNVFDSS-SAGFATG-AT----- 730  
 Db 563 EVQGNRRGANATQNLKYGAGIGTFENDRVRLRGSGSPFGLOEGVATGLATVNGAGN 622  
 OY 731 -----GLDIAINCYEGSIND-----PTSSGFTINVT 759  
 Db 623 DSAEKMGQLADLVKILGLAGLNRDFTDHSKPYTGQLRYGDAAPAGYASPRETIYAS 682  
 OY 760 SHDVTYLMKRIALSNP--NDEADRIKDELAAQVYMTSGVFPFNGEEMLRKXGNDNS 818  
 Db 683 AHDQTLMDADLVKTPPANMNSAARFONLAHSLLLGGLPFVPADDELLRSKSPDTDS 742  
 OY 819 YNADAVNEFDMS-----RKAQYPPVFNYS 844  
 Db 743 YNSGDWENALDWTGHTNGFGKGLPRAEKNENANMALYRTLLGDALAKYTPADITASDHR 802  
 OY 845 GLIHLRDHPAFRKTGTANENSHLOPLNSPENTVAYELTDHVN-KDKWCIIVVYNNPKT 903  
 Db 803 ELLEVRSSSSLFRLDTAAVOQSLSFLPAPGTGYAMKLSGVSATNPYRDLVIFNGSAD 862  
 OY 904 VATIN 908  
 Db 863 AVTILN 867

RESULT 6  
 T01688  
 starch debranching enzyme, pullulanase-type - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence, revision 19-Feb-1999 #text, change 29-Oct-1999  
 C:Accession: T01688  
 R:Beatty, M.K.; Rahman, A.; Cao, H.; Woodman, W.; Lee, M.; Myers, A.M.; James, M.G.  
 Plant Physiol. 119, 255-266, 1999  
 A:Title: Purification and molecular genetic characterization of zpu1, a pullulanase-t  
 A:Reference number: 214395; MID:99097061  
 A:Accession: T01688  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-962 <BEA>  
 A:Cross-references: EMBL:AF080567; NID:93411265; PIDN:AADJ1599.1; PID:93411266  
 A:Experimental source: endosperm  
 A:Genetics:  
 A:Gene: zpu1  
 A:Function:  
 A:Description: hydrolyzes alpha 1,6 glucosyl linkages

Query Match 15.88: Score 797; DB 2; Length 962;  
 Best Local Similarity 27.0%; Pred. No. 1.1e-37;  
 Matches 262; Conservative 143; Mismatches 325; Indels 240; Gaps 38;

OY 134 AAKPAVSNAVLADSNQVLYKLSQPLTLGEGXSGFTVHDDTANKDIPVTSVKDASLQDVT 193  
 Db 56 APPRPATAEAGASSDVG-----AAESAQGLL-----DARAYVTK 94  
 OY 194 AVLAGTFQHLIFGSDW-APDNHSTLLKKTNNLYOFSGDLPEGNYQYKVALDMSNNSTP 252  
 Db 95 SLIA-----WNISQKTSILFLYASRNATMCMSSQDMSGYSKVELQEP-NGCLP 142  
 OY 253 SDNINLTVPAGCAHYT--SYISTHAYVDTINPNADLOVESGVKTDLYVTIL--GED 307  
 Db 143 SS-----VYOKFFPIS--YRAFRIPS-VVAVTLVKQLAVASFDAGNR 185  
 OY 308 PDSHTLSITQDGYQKQVIRVNLVNSQYVSGDDLGNITTYOKATTFKVPASTPOVNV 367







```

Db 194 GWSVTGPREMERNRYLLEVDVYHPTKAQVLCACLDAPYARSLANGARTWLVDTNNETL 253
QY 441 DPAGNSDKHITPK--NIEDEYTEMADVPDSI-DPNSGMKNKGYLALTEKGTGPDNV 497
Db 254 KPASWDELADEKPKLDSFSDITIELHIDFSAHIDVTDSRGGFRAFAVQASAGMEHL 313
QY 498 KTGIDSLKQGLGTHQVLMVPFASNSVD-----E 525
Db 314 R---KLSDAQTTHVHLLPSFHFAGVDIKSNMKEFVDECELATPFGSDMOQAAYVAIOE 369
QY 526 TDPDNNMGYDPRNYDEVEGQATNANGNARIKEKEWMLSHREHIGVNMDDVYNNHFA 585
Db 370 EDPY--MNGYNPVLKMGVKGKGSASDPDGRSRIIEYRQVQALNRIGLRYMDVYNNHDS 427
QY 586 TQ---ISDFKIVEYYIYRTMIQVLIPTDQYLEMKLAEKRPVOKFTIDSLEKYVNEH 641
Db 428 SGPCGISVSLDKIYVGYVYRDTNGQIENSAMN--NTASEHMYVRLIYDOLLNNAVYK 486
QY 642 IDGPRFDLMLLGRKTM--SKAA-----SELHAINPG-TALVGEPTGGSALPDQQLTK 694
Db 487 VDGEFEDLMGHIMKRTMYTKSALQSLTJDAGVDSKITYLGEQV--DFAEVARNQGIN 544
QY 695 GAO---KMGVAVFNDNLNRLD--GNVFDSS--AOGEFATG-----A 729
Db 545 GSQLMSTGTGSGFNDRIKDALINGNPFNGPLQOGFNTGLFEPNGFTQGNADTRRSIA 604
QY 730 TGLDALKNGVEGSIND-----FTSSPGETINVTYSHDY 764
Db 605 T-YAQDQIGLAGNLRDYVLISHTEAKKGEIHTEFDGLPVGTATPIETINVSASHNE 663
QY 765 TLMKXIALSNRPSDEAD-RIKMDELAQAVMTSGVPPMGGSEEMLRKKGNDNSYNACD 823
Db 664 TLFEDIVKTPMILSVDERCRINHLASSMMSLQGIPEFHAGDELIRKSIDRDSYNSGD 723
QY 824 AVNEFDMS-----RKAQYDQVENVYSGLIHL 849
Db 724 WFNKIDFYEINNMGVGLPPSEKKNEDNMPMKPRELNSFKAKAKHIIIALDSEVDILKI 783
QY 850 RLDPHAFMTTANETINSHLOFLNSPENTVAYVELTDHVNKD-----MGNIIYV 897
Db 784 RYSSPLFRSLTANDIKQVRFRHNTGSLVPVGIYVINGIEDARSEPMAQOLDTFSYVTV 843
QY 898 YN--PNK-----YVATINLPSGKAIAINATSGKVEGSEILGQAEQVOPGISMAT 944
Db 844 FNVGCEHVSMDIPALASMGFELHPVQVNSSDTLVRSAYEATGRFTVPGRTVSV 898

RESULT 10
S51324
PulIulanase - splinach
C:Species: Splinacia oleracea (splinach)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Sep-1997
C:Accession: S51324
R:Renz, A.; Schmid, R.; Kosmann, J.; Beck, E.
submitted to the EMBL Data Library, January 1995
A:Description: Sequence of a cDNA of the Splinacia oleracea L. starch debranching enzyme.
A:Reference number: S51324
A:Accession: S51324
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-964 <Renz>
A:Cross-References: EMBL:X83969; NID:9634092; PTD:9634093

Query Match 15.2%; Score 771; DB 2; Length 964;
Best Local Similarity 26.0%; Pred. No. 3.5e-36;
Matches 260; Conservative 133; Mismatches 320; Indels 286; Gaps 39;

QY 136 LAGTQHIFGSGDWAADN-----HSTLLKRYTN----- 223
Db 10 LASSFHHTYPMRLPFLPENNFTTKLPVNSFAIGSSRSFHSPLKDDSCPCSCMAVEV 69
QY 224 -----NLYQFSGLPPEGNTQY-----KVALNDSNNNSIP 252

```

```

Db 70 GSASSVSOSELOGSLNSCRAYWPSKYTFAMNVDJNGSYYLEFASKYTAALKFTDAGIEGY- 128
QY 253 SONINLTVAGG--AHVT--FSYISTHAYVDITINNPNADLOVESGVKDTLVTYTLGEPP 308
Db 129 DVKIKLDRDQGLPANTYTERKFPHING-----YSAFAP-ATLVYDSLKQLVLAAPASAG 183
QY 309 DVSHTLSTQTDGVOAKQVILPNNVLSOYYEGDDLGNTYTKATTFKVMAPTSTOVNVL 368
Db 184 ACRNATGLQLPG-----VIDELXSYDG-PLGAVFSENTISLYLMAPTAQAVSAS 231
QY 369 LYDSATGSTKIVPMTASGHWGWEATVNQNLNMYIMEYVT--GGGSTRT-----ADPYA 422
Db 232 IFKDSGGEPLQTVOLIESNGVSAVGPRTWEGCYVVEITYVHHSTLIERKSFADPYA 291
QY 423 TIAVNGRTGMIIVDLAK--TDPAGNMSDKHITPKNIDE-----VITEMOVRDS 470
Db 292 KGISADYKRTLLADLSSEYLKPEGM-----ENLADKKPULSPDSISLLEYLHIDPS 343
QY 471 -----IDPNSGMKNKGYLALTEKGTGPDNVKGTIDSLKQGLTHVOLMPVFNASNSVD- 524
Db 344 AYDLATVHPDL-----RGGYLAFTSDSAGVNH--LEKLSAAGLTHVHLLPSEFOFAEYVD 395
QY 525 -----ETDPTQD-----MNCYDPRNVDVEGQYATNANGNA 555
Db 396 DKKKKFFVDTKRFFETLPDSEBOAQITAIREDGCMNGYNPVLWGTGPKGYATDPNCPG 455
QY 556 RIKERKEMVLSHREHIGVNMDDVYNNHFAQISD----FDTIVEYYIYRTMIQVLIPTD 611
Db 456 KIIIFRKNVQALNRIGLRVLYLVYVYNNHLSGSPSDNSVLDIVPGYILR-----DND 509
QY 612 QVLENR-----LXARPMVOKFTIDSLEKYVNEHYHIDGFRPDLMLLGRKTSKASELH 666
Db 510 GALENSTCVNDTASEHFHFWERLIDDLKHMAYNKVADGFRPDLMLHIMKHTVYKATNMQ 569
QY 667 AIN-----PGIALVGEPTGTSALPDQQLTKGAQ---KMGVAVFNDNLNRLA-LD 714
Db 570 GLSKNIDGVEGSSILYLYGEGMDFGEVA--NNAQVNASQLNMGTCIGSFNDRIRDAVIG 627
QY 715 GNVFDSS--AOGEFATGNT-----GLTDAIKNGVEGSIND----- 746
Db 628 GCPFGPPIQOQVYVYTGISLQPDNDHDSKANADRMILAVAKDHQVAGNGLRLYILTNCDS 687
QY 747 -----FTSSPGETINVTYSHDNTYLMKXIALSNPNDSEAD-RIKMDEL 788
Db 688 KOVKGEVYTYTGTPVGYAMQPIETINVSANDNEIPLDVLKPTTYITVDERCRVNL 747
QY 789 AQAVMYTSQGVPPMGGSEEMLRKKGNDNSYNAGDAVNEFDMS----- 831
Db 748 AVSIALSOGIPEFHAGDELIRKSLDRDSYNSGDMFNLDFSYSNNMGVGLPPKDNE 807
QY 832 -----RKAQYPD-----VFNYSGLIHLRLDPHAFMTTANETINSHLOFLNSP 874
Db 808 SNMPLIKRKLNPSTKPKDKNHIIAAVENTNLLQIRYSSPLRLSAKOIEDVRFRHNV 867
QY 875 EN---TVAYVELTD-HVN-----KDKWNTIYVYVNPRTYATVNLPSGK-----WAI 916
Db 868 PSWIPGLIAMSIEDHAGAPGLSQIDPKQYIVLIINQPTETKRVNPDLRKKSILQHLHY 927
QY 917 NATSGK--VGESTLQQAEGSVOPPGISMMILHQEVSPOH 953
Db 928 OSTSGDVTWKEKYEPTSCFTIIPKSTAVF---VEPRH 963

RESULT 11
A26879
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) precursor - Klebsiella pneumonia
N:Alternate names: pullulan 6-glucanohydrolase; pullulanase
C:Species: Klebsiella pneumoniae
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 06-Dec-1996
C:Accession: A26879; S02472
R:Katsura, N.; Takizawa, N.; Murooka, Y.
J. Bacteriol. 169, 2301-2306, 1987

```

A:Title: Entire nucleotide sequence of the pullulanase gene of *Klebsiella aerogenes* W70  
 A:Reference number: A26879; MUID:87194626  
 A:Note: K. aerogenes  
 A:Accession: A26879  
 A:Molecule type: DNA  
 A:Residues: 1-1096 <KAT>  
 A:Experimental source: strain W70  
 R:Charalambous, B.M.; Keen, J.N.; McPherson, M.J.  
 EMBL J. 7, 2903-2909, 1988  
 A:Title: Collagen-like sequences stabilize homotimers of a bacterial hydrolase.  
 A:Reference number: S02472; MUID:89030658  
 A:Accession: S02472  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-74 <CHA>  
 A:Note: part of this sequence was confirmed by protein sequencing  
 C:Genetics:  
 A:Gene: *pula*  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
 F:1-19/Domain: signal sequence status predicted <Sig>  
 F:20-1096/Product: alpha-dextrin endo-1,6-alpha-glucosidase status predicted <MAT>

Query Match 15 0%; Score 759.5; DB 2; Length 1096;  
 Best Local Similarity 24.5%; Pred. No. 1.9e-35;  
 Matches 279; Conservative 147; Mismatches 380; Indels 335; Gaps 46;

QY 22 VKGSAHADNNTTIIYVFCPAG-----DYQWSLMMPKDGGAEDFNQPADSFG 74  
 DB 61 VPEAVQASARQ--VHLVDINGITSTPADATKLYLMN-----NETCALSL 108  
 QY 75 AVASADIEGPNPOYGIYTRQDM---TKDVSADRYTDLKSGNEVWLEGNSTQIYNEKD 130  
 DB 109 APVA-----DWNDSYTPPTGSPDKY-----GPIYVVI----- 133  
 QY 131 AEDAAKPAVSNAVLDASNOYLVKLSOPLTLEGXSGFTVDDPANKRIPYTSVKDASLGQ 190  
 DB 134 -----PLTKESGINIVYDGT-NKLIDSGRVSESDFTD 166  
 QY 191 DWTAVALG-----TFQHIQF-----GSDWAPNHNSTL-----KKVTNNLVQFSG 230  
 DB 167 RVSIVNGSNAVYDSRADAFRAFGVALADAHV--DKTLLMPGCGENKRIYRLYYSHSS 224  
 QY 231 DLEBGNQYVALANDSNNNYPSDNTLVYPAGCAHVTSEYITSTH-AVYDTINNPNADL 289  
 DB 225 -----KVAADS--NGEFSQKYKLTPTVYNOQVSMRF---PHLASYPARKLPD-DV 269  
 QY 290 QVESGVKTDLVYTLGEDPVSHT---LSIQDGYQAKOYIPRNVLMSQYYSGDLGN 346  
 DB 270 NVDE-----LLQDGGIAESDGLSLSHFGADRRA-GRILCRRAELSTGAOL-- 318  
 QY 347 TYQKATTFKVMAPTSTQVNVLLYDSATGVTKIVPMTA--SGHQWVEATVNONLENNYVM 405  
 DB 319 --TDSGVTFVMAPTAQOVELVIT--SADKKVLSHMPTRBSASGAWMGSGDLKGFYR 375  
 QY 406 YEVTTGGGSTTAV-----DPIYATLAPNCTRGKNIYDL--AKTDPAGWNS---DKITTP 433  
 DB 376 YAMTVYVHPOSRRKVEQYEVTPVAHSLSTNSEYQOVVLDLNSALKPECDMLTTPHAKTK 435  
 QY 454 KNTDEVIYEMOYRDFSI--DPSNGMKKGYLLTEKGTGPDNVKIGDISLQOLGITHV 512  
 DB 436 ADLAKRTIHSIHRDLSAMQIYPAELRGKYLALTAO---ESNMVOHLKQLSASGVTHI 491  
 QY 513 QLMPPVANSVDE--TDPYDQ-----NNGYDPRNDVDEGOYATNANGNARIKEKKEWLSHRE 570  
 DB 492 ELIPVFDLATVNEFSDKVDADIQPFSLCEVNSAVKSSSEFAGYCDSSYEEVLTOLKON 551  
 QY 532 -----NNGYDPRNDVDEGOYATNANGNARIKEKKEWLSHRE 570  
 DB 552 DSKDNQVOVALNTLVACTDSYNNMGYDPFHYVPEGSYATDPECTARIKEPRTMIOAI-KQ 610  
 QY 571 HICVNV--MDVVYVHTFATQISD---FDKIVPEYYR-----TMIOVITPTQVLEMKL 618

DB 611 DLGMNVIMDYVYVHTNNAAGPTDRTSVLDKIVPYYORLNETGTSVESATCCSDSAPEHRM 670  
 QY 619 XAERPVOKFILOSLKXWVEYHIDGFRPDMLALGKOTSKAASELHAIPGIALYGE 678  
 DB 671 FA-----KILASLAWTTDYKIDGFRPDLMGHPAQLLSAERIKALINPDIYFFGEG 724  
 QY 679 WTGTSALPDDQLLTCKAGKGMVAFNDLRLNAL-DGNVDS-----SAQGFATGA--- 729  
 DB 725 WDSNQS--DRFEIASQINLKGCTGIFPSDRLRDAVRGCGPDSQDALRQNOVGSGAGVL 782  
 QY 730 -----TGILTD-----AINKVE-----GSINFSTSPGE 753  
 DB 783 PNELTLLSDQARHLADLTRLGAGNLADFLVLDKDAVARGSEIDYNGAPGGAADPTE 842  
 QY 754 TINVYSHDQVNTLMDKITALSNPNSEAD-RIKMDELQAQVVMYSGVFPQGGEMRXK 812  
 DB 843 VNVYVSKHNDQTLMDISTYAAQAEADLDTRVQMVAFLATVILGQCIAPDQGSSELRSK 902  
 QY 813 GGDNSYNAGDAVNEFDWS-----RRAQYPDV 839  
 DB 903 SFTFDSYDSDGDMFNRVDYSLQDNNVYNGMPRSSDGSNDIARVKDAVATPGETELKOM 962  
 QY 840 FNYYSGLIHLRDLHPARMTANEINSHLOPLN--SPENVAIYELT-----DH 885  
 DB 963 TATYQELTALRKSSPLETLDGATVMKRVDFRNTGADQDGLLVMTIDDMQAGROSGOP 1022  
 QY 886 VNKDKNIIIVVYVNPNTVATTINL-PSGKWAIVNATSGKVESTLGOAEGSVQYPGISMMI 944  
 DB 1023 CRRHRCGDQRRACKPDAAGLRHIAERYSAGCGRVVAGERVOVADSGVTLPANSAV 1082  
 QY 945 L 945  
 DB 1083 L 1083

## RESULT 12

S38801  
 alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) precursor - *Klebsiella pneumoniae*  
 N:Alternate names: pullulanase  
 C:Species: *Klebsiella pneumoniae*  
 A:Variety: ATCC 15050; strain K21  
 C:Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #ext-change 08-Oct-1999  
 A:Accession: S38801; A25025; S04864  
 R:Jansen, B.J.H.; Steyn, A.J.C.; Pretorius, I.S.  
 Curr. Genet. 24, 400-407, 1993  
 A:Title: Regional sequence homologies in starch-degrading enzymes.  
 A:Reference number: S38801; MUID:94130317  
 A:Accession: S38801  
 A:Molecule type: DNA  
 A:Residues: 1-1038 <JAN>  
 A:Cross-references: EMBL:119312  
 A:Experimental source: ATCC 15050  
 R:Chapou, C.; Balbaud, O.  
 J. Bacteriol. 164, 639-645, 1985  
 A:Title: Structure of two divergent promoters located in front of the gene encoding p  
 A:Reference number: A91807; MUID:86033621  
 A:Accession: A25025  
 A:Molecule type: DNA  
 A:Residues: 1-65 <CHA>  
 A:Cross-references: GB:M12503; NID:q149223; PIRN:AAA25087.1; PID:q551870  
 R:Kornacker, M.G.; Boyd, A.; Pugsley, A.P.; Plastow, G.S.  
 Mol. Microbiol. 3, 497-503, 1989  
 A:Title: *Klebsiella pneumoniae* strain K21: evidence for the rapid secretion of an ana  
 A:Reference number: S04864; MUID:89343636  
 A:Accession: S04864  
 A:Molecule type: protein  
 A:Residues: 'X', 21-31, 'X', 33, 'X', <KOR>  
 A:Experimental source: strain K21  
 C:Genetics:  
 A:Gene: *pula*  
 C:Keywords: glycosidase; hydrolase; lipoprotein; polysaccharide degradation  
 F:1-19/Domain: signal sequence status predicted <Sig>  
 F:20-1038/Product: pullulanase status experimental <MAT>



Db 595 KHLN---YSLFVAFYSGNERI-ETJLPKPR-KEHLAYEKIVDSTTG 635

## RESULT 14

G81717

glycosyl hydrolase family protein TC0312 (imported) - Chlamydia muridarum (strain Nigg)  
C/Species: Chlamydia muridarum, Chlamydia trachomatis Mopn  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-May-2000  
C/Accession: G81717  
R/Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gyles, M.; Nelson, W.; Debby, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A>Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A/Reference number: AB1500; M0ID:20150255  
A/Accession: G81717  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-666 <TEXT>  
A/Cross-references: GB:AE002298; GH:AE002160; MID:g7190343; PIDN:AAF39177.1; PID:g719035  
A/Experimental source: strain Nigg (Mopn)  
C/Genetics:  
A:Gene: TC0312  
C:Superfamily: glyx protein

Query Match 10.2%; Score 517.5; DB 2; Length 666;  
Best Local Similarity 26.9%; Pred. No. 5.8e-22;  
Matches 177; Conservative 105; Mismatches 244; Indels 133; Gaps 30;

355 PKVAPSTGVNVLVDASATGVTAKIVPMASCH---GVWEATVNONLENNYVYEVETGQ 411  
26 FSLVAPAAQOVLVLLDPF--SEIHEIPLSSTDRGALVHIEISGINSNSTAYKLOS 83  
412 GS-----TRTAVDPYATAI-----APNGTGMIVDLAKTDPAGNSDKAITPK 454  
84 DSAPPNSTAYIADPYSKNIFSQLGSSKQPN---DYFSYLKQDIPMEGD---TPL 137  
455 NI---EDVEYTEMVDYRDSIDPNSGKMKKRYLALTEKGTGPDNVKQIDSLKOLITTHV 512  
138 RLPENFIEIEMHVRSTQPSQVTHPGTFGLIEK-----IDHLKLGYNAV 186  
513 QLMVFNASNSVDET-DPTQDN-----WGYDPRNYDVEGOVATMANCARIKEREM 563  
187 ELVLPF---EFDETIHPKKNDFPLCLNMGYSINFCDSRRITYTDPAPAREKTL 243  
564 VLSLHREHIGVNDVYVNH-T-FATQISDFPKIYVEYYRTMIVYIIPDYLEMK----- 617  
244 VKTLHRAIGIEVLDVFNHNGEGFEGTSCPLWIDLESY---WVNNQDLNFGSGCNT 298  
618 LXAERPVAOKFIIDSLKYVNEVYHIDGFRDLNALLCKDMSKAS---ELHAIIPG--- 671  
299 VNTTPIAIIKMLDALRYWQEMHVDGFRDLAVFSDRLQGVRSLLPILQALSSDSIL 358  
672 --LALYGEPM-TGCTSALPDQDLITKGAOKMGVAVFNDLRNALDGNVDSASQGFATG 728  
359 SETMLIAEPMDAGLYQLGHPISISTMSEMG--CYRDHYKAFANDPPOVSS--FASR 414  
729 ATGLTDAIKNGVEGSIINDFTSSPGETINVTSHDNTYLMKIALSNPDSADRIKMD-- 786  
415 ISGSRDIYPAG--NSTN-----SINYICSHDGFLLDYSAVANHKNENEGENNBDGT 464  
787 -----ELAQ-----AVVMTSQGVFMQGGHEMLRKKGGNNSY 819  
465 SANSTYNGEGETTDPNICQLERQKNEFLALFLSOGIPIKISGDEYGHATVAGNNHW 524  
820 NACDANVEFDMRKRAQYPDVFNYSGLHLRLDPAFRMTTANFINSH-LQFLG----- 873  
525 CLDTKINHFLMDRLAEKKEFFSFLCOITLRTHTA--ELFNTNLSETTITWLNLSQGLR 582  
874 ---PENTVAVELTDVKNKMGNIIVYVNPKTATVINTLDSGKVAINTASGKVESTLG 929  
583 EMTDHYLAEFL-KHPN---YSLFIAFYSGNERI-ETALPLRQE-HLAVERKIVDSTTG 635

## RESULT 15

S77094

glycogen operon protein (EC 3.2.1.-) glgX-2 - Synecocystis sp. (strain PCC 6803)  
N/Alternate names: protein slr1857  
C/Species: Synecocystis sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C/Accession: S77094  
R/Kanehko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys  
s.  
A/Reference number: S74322; M0ID:97061201  
A/Accession: S77094  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-707 <RAN>  
A/Cross-references: EMBL:D90908; GB:AB001339; MID:g1652725; PIDN:BA017652.1; PID:g165  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C/Genetics:  
A:Gene: glgX-2  
C:Superfamily: glyx protein  
C/Keywords: glycosidase; hydrolase

Query Match 10.0%; Score 505.5; DB 2; Length 707;  
Best Local Similarity 26.2%; Pred. No. 3.1e-21;  
Matches 158; Conservative 89; Mismatches 223; Indels 133; Gaps 21;

345 GNTYQKATTFKVPAPSTGVNVLVDASATGVTAKIVPMAS---GVWEATVNONLEN 401  
26 GATVPGVNFSTISSTACTVTLFEKRAPQFVEIPEFSERIGNVCMVFDLDFEN 85  
402 WYVYEVETG---QC-----STRANDPYATAI-----PNGT-----RGIYDLAK 439  
86 LEVGYRMEGNNFQOGHWFDPSPKVLDPYAKVVSGRDVMTQTPWDDIYONRRL---S 141  
440 TDPAGNSDKHIIIPKNIIDEVYTEMVDYRDSIDPNSGK--NKGKYLALTEKGTGPDNV 497  
142 FDFEDWENDSPLDYV-LEDNAVIEYEMHVRGFTKDPSSVCKENHNGTFAGLISK----- 192  
498 KTGIDSLKQIGITHVQMLPVFASNSVDETPTQ-----DNMGYDPRNYDVEGOY 547  
193 ---IPYQEGVNTIELPIF---EFDEFHSRYPHETGEFLVYMGYSTVNFAPRAGY 246  
548 ATNANGNARIKEPEWYLSLHREHIGVNDVYVNH-T-FATQISDFPKIY 595  
247 AATGKFGMOIDELKNLKVLEKLVGISVLDVFNHNGEGTSCPLWIDLESY---WVNNQDLNFGSGCNT 298  
596 PEYVYRTMIOVYIIPDQVLEMKLAERPVAOKFIIDSLKYVNEVYHIDGFRDLNALLGR 655  
307 PEGYVNF-----SGTGNLNLGNRPYRGVAVLDOCLRWTAEPHIDGRFDLASILGR 358  
656 DYSKASAS---ELHAINCIA---LYGEPMVGTGTSALPDQDLITKGAOKMG--VAVEN 706  
359 DPMGYPLANPPLLETLAFDILARSKLIAEAMDAG-----GLYQVSGSPSYGRAMEWN 411  
707 DNLRNALDGVNFDSSAAGFATGATLDAIKNGVEGSIINDFTSS---PGETINVTSHDN 763  
412 GKTYD-----TVKKFKGDAGVIGEMAQRLOGSPDLYQAGARPPSINFTVAHDG 462  
764 YTLMDKIALS-----NPNDSEAD-----RIKMDLAQAVM 794  
463 FTLADLVAVNGKHNHYANGENONGANDNYSMNGVEGPTNDPILRLARANOMRAIALIL 522  
795 TSGCVPPMOGGEEMLRKKGGNDSYANGADVNEEDMSKRAQYPDVFNYSGLHLRLDHP 854  
523 VSGVPMPLMGDEMGKTQDGNNNYCHDSPPNMLMHLBONAMFRFVXICIAFRLAHP 582  
855 AFR 857

Mon Apr 23 09:45:28 2001

us-09-262-126c-2.rpr

Page 11

Db 583 VLR 585

Search completed: April 22, 2001, 09:04:38  
Job time: 419 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 09:06:01 ; Search time 52.01 Seconds

(without alignments)  
629.653 Million cell updates/sec

Title: US-09-262-126c-2

Perfect score: 5060

Sequence: 1 MAKKLIYCVLSVCLVLTWAF.....VPGISMMLHQEVSPPHGKK 956

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwisProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1570	31.0	843	1	PULA_THEMA
2	787.5	15.6	1090	1	PUL_A_KLEPN
3	759.5	15.0	1096	1	PUL_A_KLEAE
4	509	10.1	777	1	ISOA_FLASP
5	487	9.6	776	1	ISOA_PSEAY
6	487	9.6	776	1	ISOA_PSESP
7	440.5	8.7	721	1	GLGX_MYCTU
8	412	8.1	659	1	GLGX_HAEIN
9	363	7.2	657	1	GLGX_ECOLI
10	343.5	6.8	1475	1	APU_THEY
11	334	6.6	1279	1	APU_THESA
12	333.5	6.6	1481	1	APU_THETU
13	292	5.8	1861	1	APU_THETU
14	278	5.5	591	1	CDAS_BACHS
15	251.5	5.0	574	1	CDAS_THET
16	243	4.8	586	1	AMYM_BACAD
17	243	4.8	666	1	AMYB_BACCL
18	238.5	4.7	1196	1	AMYB_PABPO
19	237.5	4.7	627	1	GLGB_BACSU
20	231	4.6	585	1	NEPU_THETU
21	229	4.5	588	1	NEPU_BACST
22	225	4.4	639	1	GLGB_BACST
23	225	4.4	764	1	GLGB_STRAU
24	223	4.4	498	1	AMY3_DICTH
25	223	4.4	580	1	YF62_MYCTU
26	220	4.3	562	1	O16G_BACTR
27	220	4.3	731	1	GLGB_MYCTU
28	218.5	4.3	555	1	O16G_BACCO
29	217	4.3	520	1	AMY_BACME
30	215	4.2	703	1	CDGT_BACG2
31	202.5	4.0	639	1	GLGB_BUTRI
32	199.5	3.9	561	1	TREC_BACSU
33	198.5	3.9	919	1	AMY_STRLI

34	193.5	3.8	718	1	CDGT_BACCI	P30920 bacillus ci
35	193	3.8	770	1	GLGB_SYNY3	P52981 synechocyst
36	190.5	3.8	799	1	GLGB_MAIZE	O08047 zea mays (m
37	189	3.7	557	1	AGL_PEDPE	P43473 pedicoccus
38	187	3.7	713	1	CDGT_PABMA	P31835 paenibacill
39	185.5	3.7	718	1	CDGT_BACSS	P31747 bacillus sp
40	184.5	3.6	711	1	CDGT_BACST	P31797 bacillus st
41	184	3.6	558	1	O16G_BACCE	P21332 bacillus ce
42	181.5	3.6	718	1	CDGT_BACLI	P14014 bacillus li
43	180	3.6	713	1	CDGT_BACSP	P30921 bacillus sp
44	180	3.6	1902	1	P2P_LACPA	O02470 lactobacill
45	179.5	3.5	562	1	AMY2_DICTH	P14898 dictyoglomu

## ALIGNMENTS

RESULT ID	PUL_A_THEMA	STANDARD:	PRT:	843 AA.
AC	O33840:			
DT	30-MAY-2000 (rel. 39, last sequence update)			
DT	30-MAY-2000 (rel. 39, last annotation update)			
DE	PULULANASE PRECURSOR (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE) (PULULAN 6-GLUCANOHYDROLASE).			
GN	PUL_A OR TMI845.			
OS	Thermotoga maritima.			
OC	Bacteria; Thermotogales; Thermotoga.			
OX	NCBI_TaxID=2335;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MSB8 / DSM 3109;			
RX	MEDLINE=98115241; PubMed=9453151;			
RA	Bibel M., Bretti C., Gossler U., Kriegshaeuser G., Liebl W.,			
RT	Isolation and analysis of genes for amylolytic enzymes of the			
RL	hyperthermophilic bacterium Thermotoga maritima."			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MSB8 / DSM 3109;			
RX	MEDLINE=99287316; PubMed=10360571;			
RA	Nelson K.E., Clayton R.A., Gill S.R., Gwin M.T., Dodson R.J.,			
RA	Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,			
RA	McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,			
RA	Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,			
RA	Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,			
RA	Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;			
RT	Evidence for lateral gene transfer between Archaea and Bacteria from			
RL	genome sequence of Thermotoga maritima."			
CC	-1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYSES			
CC	(1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO			
CC	FORM MALTOSE.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO			
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: AJ001087; CAA04522.1; -			
DR	EMBL: AE001821; AAD36907.1; -			
DR	TIGR: TMI845; -			
DR	InterPro: IPR000461; -			
DR	Pfam: PF00128; alpha-amylase; 1.			
KW	Hydrolase; Glycosidase; Signal.			
FT	SIGNAL 1 19			
FT	CHAIN 20 843			
FT	PULULANASE.			



FT CONFLICT 10 10 V -> F (IN REF. 2).  
 FT CONFLICT 15 15 V -> I (IN REF. 2).  
 FT CONFLICT 23 23 G -> S (IN REF. 2).  
 FT CONFLICT 31 31 N -> S (IN REF. 2).  
 FT CONFLICT 34 34 T -> N (IN REF. 2).  
 FT CONFLICT 36 36 D -> DGNP (IN REF. 2).  
 FT CONFLICT 55 58 TAVE -> MATA (IN REF. 2).  
 SO SEQUENCE 1090 AA; 118098 MM; 240AE7DEB3FE1D6 CRC64;

Query Match 15.6%; Score 787.5; DB 1; Length 1090;  
 Best Local Similarity 25.3%; Pred. No. 1.5e-35;  
 Matches 288; Conservative 154; Mismatches 369; Indels 327; Gaps 48;

22 VKGOSAHADGNTTITIVHFCPAG-----DYOPSLMMWPKDGGAGAYDNOPADSG 74  
 49 VGEAVTAVENGA--VIHIVDIAGITSSSADYSSKNLILMN-----NEICDALS 96  
 75 AVASADIFGNPSQVGIIVFTODTKDVS-----ADRYIDLKSGNEVWLVGNSQIFYNEK 129  
 97 APVA-----DM-NDVSTTPSGSDKY-----GPYMWIPLNKE----- 126  
 130 DAEDAKKPAVSNAYLDASNOVLKLSQPLTGEKSGFTVHDDTANKDIPVTSYKMSLIG 189  
 127 -----SGCINVIYRDGTDKLDLSDLRVAFGD--FT--DRTVSVLAGNSAVYDSR-- 171  
 150 QDVAVLACTFOHIFG---GSDMAPDNHSTLL-----KVTNNLYOFGDLP---EGN 236  
 172 -----ADAFPAAGFVALAEHAWYDKN--TLLMPOGQOKPIVRLYSSSKVAADGEEK 222  
 237 YQYKVALNDSMNNSYP---SDNINTLVPAAGAHVFESYIPSTHAYVDITINPNAD--IQY 291  
 223 F-----TDKRYLKLPTPTVSQGVSMRFP---HL-----SSYAAFKLPDANAVDELLOG 266  
 232 EBGVKTDLVTYVLGDDPROVSHLSTIOFGYQAKVIRPVNLSSQVYSGDDL--GNITP 349  
 267 ET-----VAIAAEDGIIISATQVOTAG-----VLDDA-YAAEAELSTYGAQLA 309  
 350 OKATPFKVAIPSTOVNVLLYDSATGCVTKIVPMTA-SGHGWAEATVONLEMYMYWYV 408  
 310 DCGVTFRWAPTAQGVYVY--SADKKVIGSHPMFSDSAGMSQGGSDLGAFPRYAM 368  
 409 TGQGSSTFIAV-----DPAATAIPNGTRGMIVDL--AKTDPAGWNS---DKHITPKNI 456  
 369 TYVHHQSRKEVEGYEVDPAHSLSTSEYSGVVDLNDLSALKDGMNLTMPHAKQTKADL 428  
 457 EBEVLEYEMDVRFESI-DPMSCKMKGKYLALTEKKGKGDNNKTGDSLKQJCTHYVOLM 515  
 429 AKMTIHESIRLSAMDQTVPAELRGKYLAL--TAGDSNMVQHLKLTLSASGVTHVELL 484  
 516 PVFASNSYDE-TDPTQD----- 531  
 485 PVFDLATVNEFSDKADIQPFSLCEVNSAVKSSSEFACVCSGIVEVYLNDLQKSDSO 544  
 532 -----NMGYDPRNVDYBEGQVATNANGNARIKEFEKEMVLSLAREHIG 573  
 545 DNPVOALNTLVAQRTDSYMMGYDPHYTVPEGSYATDPGCTTRIKKEFRMIAI--KQDAG 603  
 574 VN--MDVYNNHTEFATQISD-----PKIPEVEYYR-----TMIOVLIPTDQVLEMLXAE 621  
 604 MNIYIDVYNNHNAAGPTDRTSYLDKIVYWTYQRLNETTGSVESATCCSDSAPEHMRFA- 662  
 622 RPNVOKFIIDSLKYVWNEYHIDGFRDLMALLGKDTMSKASLELHAINGIALYCEPWTG 681  
 663 -----KLADSLAVWTTDYKIDGFRDLNGYHPRKAQILSAMERIKRLANDIYFFGCGWS 717  
 682 GTSALPDDOLLTKGAKGKGVAVFENDLPMAL--DGNVFPSS-----SAQGFATGA----- 729  
 718 NOS--DREFIASQINLKGIGCTFSDRLDVSGVGGPFDSGALRQNGIGSAGVLPNPE 775  
 730 -----TGLDLAIKNGVEGSINDF-----TSSPGETIN 756  
 776 LASLSDQVYRHLADLTRLGMAAGNLADFMXIDKGAAGKKGSEIDYNGAPGYADPTEVVA 835

QY 757 YVTSHDNTYLMPIKIALSNPNDS-ADRIKMDLQAVVMTSGVPPMGEGEMLRKXGKN 815  
 836 YSKHNOGLMLMDISYKASQEADLATRVRMQANSLATVHMGSIARDQGSSELLRSKST 895  
 QY 816 DMSYNAGDAVNEFDMS-----RKAQYDPFVNY 842  
 896 RDSYDGDWFMFNRYDYLQDNNYVWGPRISSDGSNVEYITRVKEMVATPGAEMLKQMTAF 955  
 QY 843 YSGLIHLRLDHPAFRMTTANETNSHLOPLNSPENTAYELTQHVN-----KKNKN 893  
 956 YQELTELRKSSPLETYGDSAVKRVDFRNTGSDQAGLLVMTVDGMRKAGASLDSRLDG 1015  
 QY 894 IIVVNPNTVATINLPSGK-WAINATSGKVGESTGQ-----AEGSVQPGISMAIL 945  
 1016 LVVAIAPAESRILNPFAGETQLSLAIQGTAGENSILANGVQIADQTVLLPANSVAVL 1073

Query Match

15.0%; Score 759.5; DB 1; Length 1096;

[illegible]

```

Db 963 TAFYELTALRKSSSELTGLDGAATYMKRDEPNTGADQOTGLTWTIDDMQAGRSQSCP 1022
QY 866 VNKKDWGNIIVYVYNNKNTVATINTL-PSSKMAINATNSCWGEESTLGOAGSQVPEISMMI 944
Db 1023 CRRHGGGDRRAGKATDPAAGLRHRLPAERYASGGRPVYAGERVOYAAQGSYTLTPAMSVAV 1082
QY 945 L 945
Db 1083 L 1083

RESULT 4
ISOA_FLASP STANDARD: PRT: 777 AA.
AC 032611:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ISOAMYLASE PRECURSOR (EC 3.2.1.68).
GN IAM.
OS Flavobacterium sp.
OC Bacteria; CPB group; Flavobacteriia; Flavobacteriaceae; Flavobacterium.
OX NCBI_TaxID=229;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97340931; PubMed=9197405;
RA Krohn B.M., Barry G.F., Kishore G.M.;
RT "An isoamylase with neutral pH optimum from a Flavobacterium species: cloning, characterization and expression of the iam gene.";
RT Mol. Gen. Genet. 254:469-478(1997).
CC -1- FUNCTION: HAS A HIGH RATE OF HYDROLYSIS FOR GLYCOGEN. DOES NOT CLEAVE PULLULAN. HAS A PH OPTIMUM OF 6-7.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U90120; AAB6356.1; -.
DR InterPro: IPR000461; -.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 777 ISOAMYLASE.
FT ACT_SITE 410 410 BY SIMILARITY.
FT ACT_SITE 478 478 BY SIMILARITY.
FT ACT_SITE 533 533 BY SIMILARITY.
FT DISULFID 419 423 BY SIMILARITY.
SQ SEQUENCE 777 AA; 84340 MW; 90548988BFD1445B CRC64;

Query Match 10.1%; Score 509; DB 1; Length 777;
Best Local Similarity 25.5%; Pred. No. 1.4e-20;
Matches 191; Conservative 91; Mismatches 290; Indels 176; Gaps 27;

QY 344 LGNTV--TQKATTFVYMAPTSTQVAVLLYDSATGS--VTKIYPMYASGHWAEATV-NQN 398
Db 39 LGARDAQAQANLAFYVYSRATREVFILYKNTGSGOEVARIALSKDPATQWGLSLPST 98
QY 399 LENW-----YNYEYVG-----QGS-----TRTAVDPYATA 424
Db 99 IKNTYGITGAVYVYGRAMCPMPYDAAMTKGSAITGFVSDVDNAGRFRPNKRLDYPARE 158
QY 425 IAPNCTRCMIYD--LAKTDPAQWNSDKHI-----TPKNIEDYIYEM 464
Db 159 ISQDNTNTACDGTIYAAGAAHRRNDSGLGASKGIALAADATVSGSKPTRAKLDYIYEV 218

```







DR TIGR: H11358: -  
 DR InterPro: IPR000461: -  
 DR Pfam: PF00128: alpha-amylase: 1.  
 KW Glycogen blosynthesis: Hydrolyase: Glycosidase.  
 FT ACT\_SITE 333 333 BY SIMILARITY.  
 FT ACT\_SITE 440 440 BY SIMILARITY.  
 SO SEQUENCE 659 AA; 75290 MW; E2B71F610EBB1CAC CRC64;

Query Match 8.18; Score 412; DB 1; Length 659;  
 Best Local Similarity 24.28; Pred. No. 2.2e-15;  
 Matches 176; Conservative 110; Mismatches 246; Indels 194; Gaps 39;

338 YSSDDDLGNITTO-----KATPKVWAPSTOVNVLIDYASGTYKIVMTASGQVW 391  
 5 YNNNPIMPMSQAVENNVQITNFAISAAIGVELCFEONQETR--LEWVAT-ENVW 61  
 392 EATVNONLENNYVYEVYVQ--GSTRVAVDPYATAIAPNGRCMIVDLAKTDPAGW---- 445  
 62 HLAVTGKTEGTEVAFRIHGERFANQKILIDPYAKAV--NCK-----PDLSEESKWFLLS 115  
 446 -NSDK-HITPKNT-----EDEVIYEMVDRPS-IDPNSGKMKKXKL 484  
 116 DNRDNALAPRAVYISEEFWENDTSPNTPAETIVELHYKGFSQLNEKIPALRGTYT 175  
 485 ALTEKCTKCPNVKGTGIDSLAKOLITHVOLMPV-FASNSVD-ETDPTODMKGYDPRNYDV 542  
 176 GLAH-----PVN-----LAYLKEGLVAVELLPVNEHNEPHLARGLOANWGYPLMFA 226  
 543 PECGYATNANGNARIKEFEKVLISLREHICVNDVYVNH-----TFATQ-ISDFD 592  
 227 VERPYAAT--NNPLAEFKTMVAKFAHAGLEVILDVVFNSHASEQYPTFSQCID-- 281  
 593 KIYVEYRTMIVQYIITDOYLEMK-----LXAEPPMOKETIISDKTWNEYHIDGF 645  
 282 --QTYVWR-----NDQGRYINMTGCGNMLMSSDVGRKVVVDCLRYVEQCHIDGF 330  
 646 RFDLMALIGKDT--NSKAASELHAIN-----PGIALGEPMTGTSALPDQQLITKGAQK 698  
 331 RFDLATVLRDTPPFNSAQLFTDIKNEPSLQIKLIAEPDIO-----HYGYOV 380  
 699 CMGYAVF--NDNLRNAL-----DGNVDSASQGFATGATGLTDAIKNGVEGSIIDFT 748  
 381 GNFSYFAEWMNDRFDCRLCRFWLMKSGEL-----GAFAEFRAGSSDLFKK-----ND-- 427  
 749 SSPSETINVTYSIDNYLMDKIALS--NPNSPADRIKDE-----787  
 428 RLPRFTLNFITAHGFTLKLDELVSYNQKHNENGRDGRNENYSYHNGVEGSESLSEP 487  
 788 -----LAQAVVMTS-----QGVPMQGGEMLRXKGGNDNSYAGDAVNEFDMSRK 833  
 488 OKSAVENNRRTFASGGLMSLLLANGTMDPLLAGDEFGNTQYGNNAAYCOD--NEITWLKW 544  
 834 AOV-PDVFNYSGLIHLRLDHPAFRMTTANB--INSHLOFLN-----SPENT 877  
 545 ANFNEELFELTKOTIALRK-----QIGSLNKDDQWSDENQWMLNIGEPMTVEWMQOQT 599  
 878 VAYELTDHVKKKKGNITIVYVNPKNVATINLPSGKMAINATSGKVEESTLGAEGSVQV 937  
 600 KALOY---VIDNRW---LLLINAKAEQOMFHLPRKWKPOLIGTHNV---TLEAQOAEISS 650  
 938 PGISMM 943  
 651 MGFCHL 656

RESULT 9  
 GLGX\_ECOLI

AC GLGX\_ECOLI STANDARD: PRT: 657 AA.  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE GLYCOGEN OPERON PROTEIN GLGX (EC 3.2.1.-).  
 GN GLGX OR GLYX.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 ON NCBI\_TaxId=562;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=2978503;  
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perne N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT The complete genome sequence of Escherichia coli K-12.;  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE OF 1-590 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=89108020; PubMed=2975249;  
 RA Romeo T., Kumar A., Preiss J.;  
 RT Analysis of the Escherichia coli glycogen gene cluster suggests that  
 RT catabolic enzymes are encoded among the biosynthetic genes.;  
 RL Gene 70:363-376(1988).  
 CC -FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/  
 CC CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS.  
 CC -SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A STOP  
 CC CODON IN POSITION 507.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J01616; AAA98735.1; ALT\_SEQ.  
 DR EMBL: U18997; AAA58229.1; -  
 DR EMBL: AE000419; AAC76456.1; -  
 DR PIR: JTO400; BVBCGX.  
 DR HSSP: P10342; 1BF2.  
 DR EcoGene: EG10381; 919X.  
 DR InterPro: IPR000461; -  
 DR Pfam: PF00128; alpha-amylase: 1.  
 KW Glycogen blosynthesis: Hydrolyase: Glycosidase.  
 FT ACT\_SITE 336 336 BY SIMILARITY.  
 FT ACT\_SITE 443 443 BY SIMILARITY.  
 FT CONFLICT 288 288 Y -> YY (IN REF. 2).  
 FT SEQUENCE 657 AA; 73576 MW; AA9EFC1F67FD0420 CRC64;

Query Match 7.28; Score 363; DB 1; Length 657;  
 Best Local Similarity 23.48; Pred. No. 1e-12;  
 Matches 160; Conservative 96; Mismatches 263; Indels 166; Gaps 29;

344 LSNYTGKATTFKMAPSTOVNVLIDYASGTYKIVPMTASGQ--VNEATVNONLEN 401  
 12 LGAHYDQGVNFTLPSHAERVELCYFD-ANGQHR---YDLQHSDDIHGGLPDRAPG 67  
 402 WYIWEVETGQ-----GSTRVAVDPYATAIAPNGRCMIVDLAKTDPAGNSDKHIT 452  
 68 LRYGYRVHGPMPQAEGRHFNPAKLLIDPCARQIDGFKDNPRLHAGHNEP-DVRDNALIA 126  
 453 PKNI-----EDE-----VIYEMVDRPS-IDPNSGKMKKGLTALTEKTKG 493  
 127 PKCVVVVDHYDWDDEDAPRTPMGSTIIYEAHVGLTYLAEIPELVEIRGTAKLGH----- 181  
 494 PDNVKTGIDSLKOLGITHVOLMPV--FASNSVDEPTDODMKGYDPRNYDVPGGYATNA 551  
 182 -----PVMINILKOLGITALLELHPVADPASPRLQRMGLSNYMGVNPVAMPALPATA--C 235





Db 435 LGYDGDGMSNDFGCLDKIDKLDVLSGLISVITLNPFGSPS-----NHRVD 485  
QY 537 PRNT---DVEGQATNANGNARIKEKENVLSLHREHIGVNMVYVNHFEATQISDFK 593  
Db 486 TTDYTKIDELLGDLST-----FKKLMEADHAKGKIVLDGVFNHT-----SD-DS 529  
QY 594 IVPEYVTRMLOVLI-----PFDVLEMKLKA-----ERPVRQK--- 627  
Db 530 IYEDRYGKYLTVGLGAYQAMKQDQSKSPYGDWYEIRKPGTYEGMWGFDLPYIRQING 589  
QY 628 -----FIIDS-----LKYVYN-----EYHIDGFRFDMLLGGKDTMSKASELHAI 668  
Db 590 SEYVWVKMADPIINNPNAMSKYMWLNPGDKNVGADGHRLDVAVNADFPVHFGAINTV 649  
QY 669 NPGIALYGEPMVTGCTSALPDQDLTKGAQKGMVAVFNDLRLNLDGNTVDSQAQGFATG 728  
Db 650 KPNAPMVAEMNNDASLDLDSF-----NSVMNYLFERNVIDEILDKS---PDDG 696  
QY 729 ATGCTDAIKKQVEGSIIND-----FTSSPGETINVTSHDNTLMDKIALSN 774  
Db 697 -----NVNINPDIKAKLDQRLMSIERYPLPEVFS---TWNLGSHDTMRILTVFGYNS 747  
QY 775 PNDSADR-----IKMDELAQAVVMTSGVPEMGGEEMLRKXGDNSTYNAGDAVNEF 828  
Db 748 ADEQNSQAKADLAVKRLKLAAILQMGVPCGMSIYQDE-AGOSGKD-----PDNRRTF 801  
QY 829 DWSKKAQYPPVFNYSGLIHLRLDHPAFRMTTANEINSHLOFLSPENVAVELTJHVNK 888  
Db 802 PWGK--EDTDLQTFEKKVYVNIENENQVAK-----TGDELTETLVANDVVAFGRRIRNGK 852  
QY 889 DKMGN-----IIVVYVNPNT-----VATINLPSGKMAINATSGVGHSTL 928  
Db 853 DTEKSTSPDSVAIVYINKGAKOVSIDTTFKIRDGVAFTDLGSK-TYVQDGKI-VVEV 910  
QY 929 GQAGSVQV 937  
Db 911 GSMGAILI 919  
RESULT 11  
APU\_THESA ID APU\_THESA STANDARD; PRT: 1279 AA.  
AC P36905;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE AMYLOPULULANASE PRECURSOR (ALPHA-AMYLASE/PULULANASE) [INCLUDES:  
ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);  
PULULANASE (EC 3.2.1.41) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);  
DE (ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)].  
GN AFU.  
OS Thermomicrobacter saccharolyticum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Thermomicrobacter group; Thermomicrobacterium.  
OX NCBI\_TaxID=28896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-B6A-RI;  
RX MEDLINE=94161525; PubMed=8117096;  
RA Ramach M.V., Podkovyrov S.M., Lowe S.E., Zelkus J.G.;  
RT Cloning and sequencing of the Thermomicrobacterium saccharolyticum  
RT B6A-RI apu gene and purification and characterization of the  
RT amylopululanase from Escherichia coli";  
RT Appl. Environ. Microbiol. 60:94-101(1994).  
RN [2]  
RP IDENTIFICATION OF PROBABLE VECTOR CONTAMINATION.  
RA Robinson K.;  
RL unpublished observations (NOV-1994).  
CC -1- CATALYTIC ACTIVITY: ENDOMYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
CC -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDOLYZES  
CC (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO  
CC FORM MALTOSE.  
CC FORM MALTOSE.

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN BY THE PRESENCE OF  
CC AN EXTRA C-TERMINAL SEGMENT OF 9 RESIDUES THAT SEEMS TO ORIGINATE  
CC FROM A PUC-TYPE VECTOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).  
CC -----  
DR EMBL: L07762; AAA19800.1; ALT\_SEQ.  
DR InterPro: IPR000461; -  
DR InterPro: IPR001777; -  
DR Pfam: PF00128; alpha-amyase; 1.  
DR Pfam: PF00041; fn3; 2.  
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; signal; Repeat;  
KW Multifunctional enzyme.  
FT CHAIN 1 35  
FT ACT\_SITE 36 1279  
FT ACT\_SITE 629 629  
FT ACT\_SITE 658 658  
FT ACT\_SITE 735 735  
FT DOMAIN 929 1017  
FT DOMAIN 1156 1248  
SQ SEQUENCE 1279 AA: 142430 MW: 0950CCGCA91624DD CMC64;  
Query Match 6.6%; Score 334; DB 1; Length 1279;  
Best Local Similarity 20.0%; Pred. No. 1e-10;  
Matches 225; Conservative 142; Mismatches 379; Indels 378; Gaps 54;  
QY 3 KLLIYVCLSVCLVTMAF-NVKGQSAHADGNTTII--IYHFCPAGIYQ-----WSLW 53  
Db 8 KKFISFVMSLLVLTAAFSMPFHNVYADNASVAVNI-----GDFQQLGDSNMNI- 60  
QY 54 MMRKDG-----GGAEYFQNPADSFAGVAVASADIGNPQVGIYRTQ-DMTKQVSAD 104  
Db 61 ----DSNITTMQVYGNGLYETFTPTQLKAGSYQYKVALHSMWGGVPSGNTLTLTND 116  
QY 105 RYIDLSKGNEMVWLVGNSQJTFNEKDAEDAKPVAVYLDASNOVLVKISQPLTLEGX 164  
Db 117 SYVTF-----W-----FDYNTQSTDSFK-----YPIISMDKLR----- 146  
QY 165 SGFTVHDDTANKDIPVTSVKDASLGQDVYAVLAGTFQIHFG-GSDMAPDNHST-LKKVT 222  
Db 147 -----LVGTISQAIAGAGDMDPSTSTAIMIDNF 175  
QY 223 NNLVQFSGDPEGNQYKVALNDSMNSYPSD-----NINLVAPGAVHTSYIPSTH 276  
Db 176 DNYISTYAHIPKDDYQYKVLGNTMAENYANGVQDSNQLQSLV-ANDADITFTFYDANTH 234  
QY 277 AV-----YDTINPNADLVESGVKTDLVTVLGGEDPDVSHLSIQTDG 320  
Db 235 NIMTNSPTLTGLDNNIYYDDLKHDTHD-----PEFRNPFCAIKVQ-----TVTLRIQAKN 286  
QY 321 YQAKQVIRVAVLSSQVYGGDGLGNTYQKATTFKVMAPFSTOVNVLVDSATGVTKI 380  
Db 287 HD-----LESARIST-WDDINKRTE----- 306  
QY 381 VPMTAGS-----HGWE--ATVQNLNNYMYEVTGQSTRAVDPAVATAPNCTRG 432  
Db 307 LPMTRIGESPDGNTYEWEIKLSFDHPRIWYFLKKG--TKTAV--YQDNDQLAGLG 361  
QY 433 MYVDLAKTDPAGNMSDKHITPKNIEDEVIYEMDVRFESIDPNSGKKGKYLATKGTG 492  
Db 362 KADTVYKND-----FELAYQDNFPTNDMMKGAVAYQIFPDFFYN 401  
QY 493 G-----PDN-----VKTG 500

```

Db 402 GDTISNDIAKTLISGNDPIEFHNDMNDLDPNNAAGTPGYTGDGLMSNDFEGDLKIDDK 461
QY 501 IDLSKOLGITHVQLMPYFASNSVDEPTQDNMGYDPFNRYDVEGQYATNANGARIKEF 560
Db 462 LDYKLGAGSVIYIINPIFESPS-----NHKYDTADYIKRIDEMFETT-----QDF 505
QY 561 KEWYLSJHREHIGVNMVYVHHTATQI-----SDPKIVPEY--YR 601
Db 506 EKLMSDHAHAKINIIIDGVFNHTSDSIYFNRYGKPDLAGYODMKDQNSGLSPYGDWY- 564
QY 602 TMIOVIIPTDQVLEMKLAEK-PMVQ-----KTIIDS-----LKYWVEYH-- 641
Db 565 -----TINSDDTYECWMCYDLSLPIYKSLNGSEYNTSWANFTINDKNAISKYWLMPDENL 619
QY 642 ---IDGFEFDLMALLGKDTMSKASELHAIPGIALYGEPTGTGSAIPDQULLTKGAQK 698
Db 620 NDGADGMRLDVENEVAHDFWTHFRDAINTVPEAPMIEKN--GDASL---DLGDSFNS 674
QY 699 GNGVAVENDIRMALDGNVPDSSAQGFATGTLDAIKNGE--GSINDFTSSPG--ETI 755
Db 675 VMNQY-FRNDIIDLFLIGSFD--GNGQHNPIDAKLLQRLMSIYERYPLPAFYSTM 728
QY 756 NYVSHDNYTIMDKIALS-----NPNDSADRIKMDLGAAYVMTSQGVPMQ--GEEM 808
Db 729 NLLGSHDMSITLYFGYNSADPNENSDAKRLAEQKLKATIL-OMGYPGMADIIYGD- 786
QY 809 LRKXGNDNSYNAGDAVEFDMSRK-AQYPPVFNYSGLIHLRDPAFRMTTANEINSH 867
Db 787 AGVSGGRD-----PDDRITFPWGNEDTALQDFK-----NVSSTNN 823
QY 868 LQTLNSPENTVAYELTDHV-----NKDKWN-----ITVYVNPRTVATINLPSGK 913
Db 824 NOVLITGDLLETLYAQNDVYAIIGRLINCKDAFGNSYDPSAIIAIVNRSNSDOULTIDTJK 883
QY 914 MAINATSG---VGES--TLQAGESYVPGISMAIILHOEVSFD 952
Db 884 FLRGVAFKDLINGDKSTYINGGOTTINIPAMSGVMLISDGD 927

RESULT 12
APU_THEET
ID THEET STANDARD: PRT: 1481 AA.
AC P38939;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE AMYLOPULULINASE PRECURSOR (ALPHA-AMYLASE/PULULINASE) [INCLUDES:
DE ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);
DE PULULINASE (EC 3.2.1.41) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)
DE (ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)].
GN APU.
OS Thermomicrobacter ethanolicus (Clostridium thermohydrosulfuricum).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermomicrobacter group; Thermomicrobacter.
OX NCBI_TaxID=1757;
RN 111
RP SEQUENCE FROM N.A.
RX STRAIN=39E;
RX MEDLINE=90147689; PubMed=2302196;
RA Mathupala S.P., Saha B.C., Zeikus J.G.;
RT "Substrate competition and specificity at the active site of
RT amylopullulanase from Clostridium thermohydrosulfuricum.";
RL Biochem. Biophys. Res. Commun. 166:126-132(1990).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME. HYDROLYZES
CC (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO
CC FORM MALTOSE.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; M97665; AAA23201.1; -.
DR PIR; S28669; S28669.
DR InterPro; IPR000461; -.
DR InterPro; IPR001777; -.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF00041; fn3; 2.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;
KW Multifunctional enzyme.
FT SIGNAL 1 31
FT CHAIN 32 1481
FT ACT_SITE 628 628
FT ACT_SITE 657 657
FT ACT_SITE 734 734
FT DOMAIN 926 1015
FT DOMAIN 1158 1250
FT SEQUENCE 1481 AA; 166362 MW; CB2635960539CAID CRC64;

Query Match 6.6%; Score 333.5; DB 1; Length 1481;
Best Local Similarity 20.9%; Pred. No. 1.3e-10;
Matches 218; Conservative 156; Mismatches 339; Indels 339; Gaps 52;

QY 70 ADSEFVASADIPGN-PSQVGIIVRTQDKDSADRYIDLSKGN-----EVLVEGN 121
Db 31 AETDPAIAPLANVVGQFQSGIG--DSDW--NINDKTYMTKNGGFEFTTPVALPAD 84
QY 122 SOLFNEKDA-----EDAKPAVSNAYLDASNOVLKLSQPLTLGEGSGFTVHDTANK 176
Db 85 ---YEVYALNHSWEGGVPSQGNLSLHDSVY-----TFYVNTSSVTSTKYT 134
QY 177 DIPVTSVADASIGDQVAVLACTFOHIFG-GSDMAPDNHSTLLKVT--NLLQFSGDLP 234
Db 135 PIP-----EKLPRIVGTIOSAIGAGDMKPESTAIMRKFNRYEYTAANVP 184
QY 235 GNYQYKVALNDSWNSY-----PSDNLITVPAGAHVTFYSIPSTHAYVDTINPNNA- 287
Db 185 RYERKVTILGPEWMDINTGNGQNGPNIPLAN-AVDTKTFYDYSVSHIMTDVNPILG 243
QY 288 -----DLOVES-----GVKT-DLVTVTLGEDPDVSHLSIOTDGYQAKQVIP 328
Db 244 PDNNIYYDDLKHDTHDPFRFAFGAIKTDGYVTLRI-----QAK----- 282
QY 329 RVLNLSQYYSDDLGNGTYTQKATTFYMAPTSTOVANVLYLDSATGSTYKIVPTASG- 387
Db 283 NNDLESASIKSY-WDDIKTKRTE-----VPIKTIQ 311
QY 388 -----HGWEATVNONLEN--WYVWEVYTGOSTRAVDPAATAPNGTRGMIVDLAKT 440
Db 312 SPDGQYEWWEVWKLSPDYPRIVYFLIKDG---TKTAY--YQDNDQGLGCVKADTVNKK 366
QY 441 DPAGNNSDKHT-TPKNIDEVYEM-----DVRDFS--IDP----- 473
Db 367 DEFLTLYYKKNLDTPPMKAGAVVYQIFPDRFYNGDPLNRLKESYSGFDPVEYHDDMYDLP 426
QY 474 -NSGKNKNGKYLALDEKTKGPD-----ANKTGISLKOGLIHYVOLMPYFASNSYDE 525
Db 427 DNPNDKXDPGY---TGDGIWANDFFGGDLGINDKLDYLNKLGISVYIYNPIFOSPS--- 480
QY 526 TDPTDQNGYDPRNRYDVEGQYATNANGARIKE-----EKWYLSJHREHIGVNMVY 578
Db 481 -----NHRYDFTDY-----TKIDELLGDLDPFTKTLKEMAHNGIYVILDG 520
QY 579 VYNHFPATQI-----SDPKIVPEYTYR----- 601

```

```

Db 521 VFNHTSDSIYEDRKYGLJONELGATYQAMKQGDGSKSPYGDWMEIKRDPGYIEGMMGFDSL 560
QY 602 TMIOVITITDOVLEMKLAEPNOKFTIDS---LKYWN---EYHIDGFRFDIAL 65322GFRFDIAL
Db 581 PVIHQINSGE---YVWKSWAD-----FLINPNIAISKYTMINPGDOKDAGACGRILDVANE 63322GGRILDVANE
QY 653 LKQPTSKNASELHAININGIALYGEPMPTGGTSLPDDQLLTKGAKQKMGAAVENDNLRNA 7122
Db 653 IAHDFWYFRAIINTYKRNPMATMALMGDASLDLLGDSF-----NSVMNYLERNA 6622
QY 713 LDGWFPOSSAAGFATGATGTLTDAIKNGEVSIND-----PYSSPGQETINY 758
Db 683 VIDFILDQK---FDDG-----NVYHNPIDAKLDQRLMSTIERPLPLVFS---TMNLL 730
QY 759 TSHONTYLTMDKIALISNPNDSFADR-----TKDELAQAAVMTSGCVPRFGGCEBMLRXK 8122
Db 731 GSHDTMLRLTFVGCYNSAENONSGEANDLAVRKILKLAITLQMGYPCMPSTIYGDG---AGOS 789
QY 813 GGNDSNYACDVAVNEPFSMRKAOYPDVFYNYSGILRLDHPAFRMTAINEISHILOFLN 872
Db 790 GCKD-----PNKRTFSMGREDK---DLQDFPKAVVNTIRNEOYLK-----TGDELETTY 835
QY 873 SPENTVAVELTDHVNKKDKWCN-----IIVVYNPKVTATINPLSGKWA-----INATS 920
Db 836 ANGDUYAFGRRIITNGKDVFGNSYPDVAIVYINKGE-AKSVQIDITTEFVADGVAIFDALS 894
QY 921 GKVGESTLTGAEGESVOYFGJSM 942
Db 895 GK-----TYVWAGQIIVEYVAL 912

```

```

RESULT 13
APU_THEMEU
ID APU_THEMEU STANDARD: PRT: 1861 AA.
AC P38536;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE AMYLOPULULINASE PRECURSOR (ALPHA-AMYLASE/PULULINASE) (PULULINASE
DE TYPE II) [INCLUDES: ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN
DE GLUCANOHYDROLASE); PULULINASE (EC 3.2.1.41) (1,4-ALPHA-D-GLUCAN
DE GLUCANOHYDROLASE) (ALPHA-D-XYLO-1,6-ALPHA-GLUCOSIDASE)].
GN AMB.
OS Thermomanaerobacter thermosulfurogenes (Clostridium
OS thermosulfurogenes).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermomanaerobacter group; Thermomanaerobacterium.
OX NCBI_TaxID=33950;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 3896 / EMI;
RX MEDLINE=94252998; PubMed=8195085;
RA Matuschek M., Burchardt G., Sahm K., Bahl H.;
RT "Pullulanase of Thermomanaerobacterium thermosulfurogenes EMI
RT (Clostridium thermosulfurogenes): molecular analysis of the gene,
RT composite structure of the enzyme, and a common model for its
RT attachment to the cell surface.";
RL J. Bacteriol. 176:3295-3302(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME. HYDROLYZES
CC (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO
CC FORM MALTOSE.
CC -1- SUBCELLULAR LOCATION: CELL-BOUND. IT C-TERMINUS MAY SERVE AS AN
CC S-LAYER ANCHOR.
CC -1- PTM: GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -1- SIMILARITY: CONTAINS 2 FIBROECTIN TYPE IIT-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 S-LAYER HOMOLOG (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL: M57692; AAB00841.1; -.	
DR	InterPro: IPR000461; -.	
DR	InterPro: IPR001119; -.	
DR	InterPro: IPR001777; -.	
DR	Pfam: PF00395; SLH. 3.	
DR	Pfam: PF00128; alpha-amylase; 1.	
DR	Pfam: PF00041; fn3. 2.	
DR	PROSITE: PS01072; SLH_DOMAIN. 3.	
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;	
KW	Multifunctional enzyme; Glycoprotein.	
FT	SIGNAL	1
FT	CHAIN	36
FT	ACT_SITE	628
FT	ACT_SITE	657
FT	ACT_SITE	734
FT	DOMAIN	928
FT	DOMAIN	1157
FT	DOMAIN	1681
FT	DOMAIN	1740
FT	DOMAIN	1804
FT	DOMAIN	1734
SO	SEQUENCE	1861 AA; 206104 MW; 06c23070e43b574 CRC64;

Query Match 5.8% Score 292; DB 1; Length 1861;  
Best Local Similarity 21.0% Pred. NO. 3.4e-08;  
Matches 208; Conservative 123; Mismatches 352; Indels 308; Gaps 47;

OY		178	I P V T S V D A S L G O D V T V I A L G F P H I E G S S W M A R N D S T L K K Y T N L X Q P S --- C D J P E C	233
Dd		28	I P L S V I A A D - M A S V A N I I G E F P O D U G D S N W M I D S I T T I M O Y W G N L I E F F T P T O L K A G	86
OY		236	N Y O Q V A L N D S W M N S Y P S D I N I L T V P - A G S A H Y F S Y I P S T H A Y Y D - T I N N P N A D L O V E S	293
Dd		87	S Y Q K A L L N S M S D G G V P S G C G N L L N L A N D S Y T V F W E D Y T N Q S Y T B S T K T T P L A N D K L P R	146
OY		294	G V K T D L T V T I L G E D - P D V S H I T L S I Q T D G Y Q A K O V I P R N V L N S S Q Y Y S C D ---- D L G N	346
Dd		147	L V G T I O S A I G A G N D M K P E T S - T A I M T D N E D ----- N Y S T A H A P K G D Y O K Y K T L G N	198
OY		347	T Y T O K ----- A T T P A V W --- A P T S Q V - N V L L Y D -----	371
Dd		199	T W D E N T A N G A Y K D G S N I Q I N V T N D A D I T T Y Y D A N T I N I M N Y S I L T G L D N N I Y T D L K H	258
OY		372	----- S A T G S Y T K I --- V P M T A S G H C ---	389
Dd		259	D T H D S F P R N P G A K V Q D Y T V L T R I O A K N H D L E S A R I S Y M W D I K I R I E L P M T R I G E S P D C	318
OY		390	---VW-E--A T V N O N L E M W Y M Y E V T G G S T R T A D P Y A T A I A P N G T R G M I V D L A K T D P A G	444
Dd		319	N F E Y W E I K L S F D H P T R I M Y F I L K D G ---T K T A V - Y G D N D D O L G G V K A K T D Y V N K D ---	370
OY		445	M W S D K H I T T P N I E B E V I E M D V D P F S I D P R S K K N K K Y A L I L E K G K G -----	493
Dd		371	-----F E L T Y I D K N F D T P D M K G A V Y O I F P D R E Y N G G T S N D H A K T L S	413
OY		494	-----P D N -----V K T O I D S L K O L G I T H Y Q	513
Dd		414	R G N D P I E F H K N M W D L P D N P N N A C T P G Y T G G I G S I N S D P F G L K G I D K L D Y L I K G L G V S Y I T	473
OY		514	L M P F A S N S V D E T P D O D N G M Y D P R N Y D V B E G O Y A T N A N G M N A R I K E K E K E V A L S L H R E H I	573
Dd		474	L N P I F E S P S -----N H K Y D T A D Y T K I D E M G C T T -----O P E K I L M S D A H K A G I K	517
OY		574	V M N D V Y N H F A T O I S P F D K I V E P Y -----Y T M I O V I I P -----T D Q V L	614
Dd		518	I I I D G V E N H I ---S D -D S I Y E N R R G K Y P G L G A Y Q A M K G N O S L S P Y G M Y I N S G I Y	571

OY 615 EMKLAER-PMVQ-----KEIDS---LKYWN-----EYHIDGFEFDLMA 651  
 DB 572 ECMMGYDLSLPYKISLNGSEYVNTSWANFIINDENAIISKYMLNDGADGMRDLYEN 631  
 OY 652 LUGADTKSKASELHAINPGIALYCEPWTGTSALPDQOLLTKGAKGKMGVAVFNDNLRN 711  
 DB 632 EVADHFWHFNRNAINTVFEAPMAIENM--GDASL---DLGDSFNSVWNO--FRNDIID 685  
 OY 712 ALDGNVFDSSAQGFATGATGLDAIKNGVE--GSINDFTSSPG--ETINYVTHDYTLMD 768  
 DB 686 FLIGOSFDD-----GNGCHNPIDAKKLDORLMSIERYPLPAFYSTMLILGSHDMRLIT 740  
 OY 769 KIALS---NPNDEADIKKDELQAAVYMTSOGVPFMQG---GEEMLRXKGNDSYNA 821  
 DB 741 VEGYNSADPNENSDAKQLAQKCLKLATIL-QMGYPGMADITYGDE-AGVSGKX-----793  
 OY 822 GDVAWNEPMSRK-AQPPVFNYSGLHLRLDHPAFRMTTANELNSHLQFLNSPENTYAY 880  
 DB 794 PDDRRTPFMGNEFTTLQDFKNIIS-----SINNNQVLKTGDLLETLY 835  
 OY 881 ELTDHV-----NKDKWGN-----IIVYVNPRTVATINLPSGKAINATSGK--VG 924  
 DB 836 AONDVYALGRRIINGKDKAFGTSYDPSAAIVAINRSKSDKQALDITTKFLRDGVTFKDLIN 895  
 OY 925 ES---TLQAEQSVQVPCISMMILHQEVSPD 952  
 DB 896 NNVSYSISNGOIVIDVPMASGVMILSDGQD 926

## RESULT 14

CDAS\_BACSH STANDARD; PRT: 591 AA.

AC 008341;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CYCLOMALTODEXTRINASE (EC 3.2.1.54) (CDASE) (CYCLOMALTODEXTRIN  
 DE HYDROLASE, DECYCLING).  
 OS Bacillus sphaericus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1421;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-E-244;  
 RX MEDLINE=93312539; PubMed=7763728;  
 RA Oguma T., Matsuyama A., Kikuchi M., Nakano E.;  
 RT "Cloning and sequence analysis of the cyclomaltoextrinase gene from  
 RT Bacillus sphaericus and expression in Escherichia coli cells.";  
 RL Appl. Microbiol. Biotechnol. 39:197-203(1993).  
 CC -1- FUNCTION: HYDROLYSES CYCLODEXTRINS. CAN ALSO ACT ON LINEAR  
 CC MALTODEXTRINS, WITH THE EXCEPTION OF MALTOSE. THE PH OPTIMUM IS  
 CC PH 8.0 AND THE OPTIMAL TEMPERATURE IS 45 DEGREES C.  
 CC -1- CATALYTIC ACTIVITY: CYCLOMALTODEXTRIN + H(2)O = LINEAR  
 CC MALTODEXTRIN.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- INDUCTION: BY CYCLODEXTRINS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY. CDASE MOST CLOSELY RESEMBLES  
 CC NUCOPULITANASE FROM B. STEROTHERMOPHILUS AND ALPHA-AMYLASE-  
 CC PULULITANASE FROM C. THERMOPHILUSULFURICUM.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X62576; CAA44454.1; -;  
 DR InterPro; IPR000461; -;

DR Pfam; PF00128; alpha-amylase; 1.  
 KW Hydrolyase; glycosidase.  
 FT ACT\_SITE 327 BY SIMILARITY.  
 FT ACT\_SITE 356 BY SIMILARITY.  
 FT ACT\_SITE 423 BY SIMILARITY.  
 FT ACT\_SITE 423 BY SIMILARITY.  
 SO SEQUENCE 591 AA; 67899 MW; 702FD/BA99A23E3F CRC64;

Query Match 5.5%; Score 278; DB 1; Length 591;  
 Best Local Similarity 24.5%; Pred. No. 3.8e-08;  
 Matches 135; Conservative 75; Mismatches 207; Indels 134; Gaps 29;

OY 448 DKHITPKNTEDEVIEYEMDRP-----SIDPNSKKNKRYALTEKGTGPDNVTG-- 500  
 DB 123 DVHTPPAWYKEALFYQIFERFANGDPSNDP-EGVQEMG-----GRPSAGNFFGGL 173  
 OY 501 -----IDSLKDLGITTHQIMPFVFNASNSVDELDPDQDNKGYPRNYDVEGGYATNANGN 554  
 DB 174 OGVIDHLDYLSLQGVNLFNPLFAATY-----NHRKYPTADYMKIDPQFGTN---- 220  
 OY 555 ARIKEFEKVVLSLHREHIGVNNDVYNN--HTFATQISDFDKIYPEY-----YRTMIO 605  
 DB 221 ---EKULKELVDACHARGMVLLDAVFNHCGHTFPFVDVLANGLNSRYADMFRHREMPLR 277  
 OY 606 VI--IPTQVLEH-----KLAEPRPVOKFIIDSILKYVNEYHIDGFRFDLALLGKDTM 658  
 DB 278 VVDGIPTDYDTAFEPIMPKLNTGNEVKAYLLNVGRYWEEMGLDQMRLDVANEVDHQFM 337  
 OY 659 SKAASELHAINPGIALYGE-----PWTGTSALPDQOLLTGAOKGKMGVAVFNDNLRNA 712  
 DB 338 REFRESEIKRINPSAVYILGIMHDSMPWLOG-----DQF-----DAWNNPPTNI 381  
 OY 713 LDGNVFDSSAQGFA---TGATGLTDAIKNGVEGSIINDFTSSPGETINYVTHDYTLMDK 769  
 DB 382 L-----LNFARRLTNAEFAQIIGTQAGYPOQVTE---VSENLGSHDITRL--- 427  
 OY 770 IALSNPNDSEADRIKMDLQAAVMTSOGVPPMGGEEMLRRKKGNDNSYNAGDAVNE-- 827  
 DB 428 LTLGSGN---VERMKLATLFO--LTYGCTPCITYGDEI-----GMDGEY---DPLNRKC 473  
 OY 828 FWSKRAQVDPVFNYSGLHLRLDHPAFRMTTANELNSHLQFLNSPENTYAYELTH 885  
 DB 474 MEMDRSKONTPELLAFERMSISLRKAHPALR-----GSLRFLPLVLEHPOLLYVERMW- 525  
 OY 886 VNKDKMGNIIYYVN---PKRTVATINLPSGKM-AINATSGKGESTLQAEQSVQ--VPG 939  
 DB 526 ---DNERFLIMLNNEAPVNVYIPAAQPGASMRVYV-----GECVAVEESSIALALP 576  
 OY 940 ISMTILHQEVS 950  
 DB 577 YGYALIHAPIA 587

## RESULT 15

CDAS\_THERT STANDARD; PRT: 574 AA.

AC P29964;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE CYCLOMALTODEXTRINASE (EC 3.2.1.54) (CDASE) (CYCLOMALTODEXTRIN  
 DE HYDROLASE, DECYCLING).  
 OS Thermomonasobacter ethanolicus (Clostridium thermohydrosulfuricum).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Thermomonasobacter group; Thermomonasobacter.  
 OX NCBI\_TaxID=1157;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.  
 RC STRAIN-39E;  
 RX MEDLINE=92355516; PubMed=1644767;  
 RA Podkovyrov S.M., Zelkus J.G.;  
 RT "Structure of the gene encoding cyclomaltoextrinase from Clostridium  
 RT thermohydrosulfuricum 39E and characterization of the enzyme purified

RT from Escherichia coli.  
 RL J. Bacteriol. 174:5400-5405(1992).  
 RP ACTIVE SITES.  
 RX MEDLINE-9316171; PubMed-8425614;  
 RA Podkovyrov S.M., Burdette D., Zeikus J.G.;  
 RT "Analysis of the catalytic center of cyclomaltodextrinase from  
 Thermobacter ethanolicus 39E";  
 CC FEBS Lett. 317:259-262(1993).  
 CC -FUNCTION: HYDROLYSES CYCLODEXTRINS. CAN ALSO ACT ON LINEAR  
 CC MALTODEXTRINS, WITH THE EXCEPTION OF MALTOSE.  
 CC -CATALYTIC ACTIVITY: CYCLOMALTODEXTRIN + H(2)O = LINEAR  
 CC MALTODEXTRIN.  
 CC -SUBUNIT: MONOMER.  
 CC -SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY. CDASE MOST CLOSELY RESEMBLES  
 CC NEOPULULANASE FROM B. STEAROTHERMOPHILUS AND ALPHA-AMYLASE-  
 CC PULULANASE FROM C. THERMOPHILUS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, M88602; AAA23219.1; -.  
 CC DR PIR; A42950; A42950.  
 CC DR InterPro: IPR000461; -.  
 CC DR Pfam: PF00128; alpha-amylase; 1.  
 CC KM Hydrolyase; Glycosidase.  
 CC FT ACT\_SITE 325 325  
 CC FT ACT\_SITE 354 354  
 CC FT ACT\_SITE 421 421  
 CC FT MUTAGEN 325 325  
 CC FT MUTAGEN 354 354  
 CC FT MUTAGEN 421 421  
 CC FT CONFLICT 4 4 E -> G (IN AA SEQUENCE).  
 CC SEQUENCE 574 AA: 68026 MW: E7E478D25AD6D0A CRC64;  
 SO  
 Query Match 5.0%; Score 251.5; DB 1; Length 574;  
 Best local similarity 22.9%; Pred. No. 1e-06;  
 Matches 139; Conservative 74; Mismatches 212; Indels 181; Gaps 28;  
 OY 379 KIVPM-----TASGHVGEATVNQLENVYMYEVTGQSTRTAVPYATAIAPNGTRGMI 434  
 DB 52 KIRPMVLTHNELLEDYETLELNKKFVFFYVSDGKEL-----YETE----- 96  
 OY 435 VDLAKTDPAQ--W-----NSDKHITPKNTEDEVYEMDVADPSIDPNSGKMKKGY 483  
 DB 97 AGFYKKRPNHFWGFHYPIYIGKDFVFAPEMVSQCNVQIIPERF-----NNGDK----- 147  
 OY 484 LALTEKSTGSPDNVY-----TGIDSLKQIGITHVOLMPVAFASNV 523  
 DB 148 -----SNDEPVKKPWEKPTADSFGDLOGIIDKIDYLDKDGAINAIYLTPIFLSHST 200  
 OY 524 DETDPTODNMGYDRNRYDVEGQATNANGNARIKEFKEMVLSLREHIGVNMDDVYVYNT 583  
 DB 201 HKYD-TTDYTYIDPHFGDT-----QKARELVQKCHDNGIKVIFDAVENHC 244  
 OY 584 -----PATQISDPKIVPEYVYRMIOVI---IP-----DOVLEM-KLXAERPMV 625  
 DB 245 GYDFEAFQ--DVTKNGKSKYMDPFNIYEMPIITHGKPSYEADVYWRMPKMLTKNPEV 302  
 OY 626 QKEFIISLKYVNEEYIIGDFRDLMLLGGKDTSKASELHAIINPGIATLGEPTGTSA 685  
 DB 303 QKYLLEVAEYKIKEDIDGKRLDVANEIDHFRKRFREYVKAARPAIIVGEVWHDAFSPW 362  
 OY 686 LPDDQLITKGAQKMGVAVFNDLRLNALDGNVFDSSAQGFATGA--TGLTDAIKNGVEG 742  
 DB 363 LRGDQF-----DSVANYPFRNA-----YVDFAKRKISASRFNTMITEDLMRMD- 407

OY 743 SINDETSSPETIINYVTSHDNVYTLMDKIALSNPNDSEADRIKMDLDAQAVYMTSGVPEFM 802  
 DB 408 SVNRV-----MFWLIGSHDTERF---LTLAN-----GVYARPKLALVQFTFVGIPYI 452  
 OY 803 QGGEEMLRXKGGNDSYNAGDAVNEFD-----WSRKQYDPVFNRYSGLIHLRLDHP 854  
 DB 453 YGDEV-----GWDGDDPDCRCRCKIWEKEKQNSIFNFYKLLISIRE-- 496  
 OY 855 AFRWTTANEINSHLOF-----LNSPENTVAVELTDHVNKKDKWGNIIYVYN--PKTVATY 907  
 DB 497 -----NEELKYSFCTIYVIGVFAFK-----REYKKSIIYVVLNNSNQEVIFL 541  
 OY 908 NLPSGK 913  
 DB 542 NEVEGK 547

Search completed: April 22, 2001, 09:11:53  
 Job time: 352 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 09:04:59 ; Search time 74.41 Seconds

(without alignments)  
1505.856 Million cell updates/sec

Title: US-09-262-126c-2

Perfect score: 5060  
Sequence: 1 MAAKLIYCLSVCLVLTMAF.....VPGISMMLHQEVSPDHGK 956

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1428.5	28.2	849	2	Q9XDB5
2	1252	24.7	668	2	P71095
3	1137.5	22.5	825	2	Q59319
4	1130	22.3	718	2	Q69008
5	1084.5	21.4	717	2	Q9K705
6	1038.5	20.5	718	2	Q34587
7	890.5	17.6	1798	2	Q9K211
8	819	16.2	1072	2	Q9K6N1
9	817.5	16.2	1938	2	P70983
10	815.5	16.1	910	2	Q9RXB0
11	805	15.9	988	10	Q64454
12	797	15.8	962	10	Q81638
13	793	15.7	986	10	P93416
14	787	15.6	962	10	Q93758
15	786.5	15.5	904	10	Q48541
16	771	15.2	964	10	Q41386
17	533.5	10.5	666	2	Q84046
18	531	10.5	762	2	Q9RM63
19	517.5	10.2	666	2	Q9PK26

20	508	10.0	702	10	Q9M055	Q9M055 arabidopsis
21	506	10.0	783	10	Q04196	Q04196 arabidopsis
22	505.5	10.0	707	2	P73608	P73608 synechocyst
23	497.5	9.8	718	1	P95868	P95868 sulfolobus
24	491.5	9.7	733	10	Q80403	Q80403 oryza sativ
25	480.5	9.5	789	10	Q22637	Q22637 zea mays (m
26	474.5	9.4	823	2	Q9LC80	Q9LC80 streptomyce
27	472.5	9.3	664	2	Q928P5	Q928P5 chlamydia p
28	471.5	9.3	720	2	Q9RXP5	Q9RXP5 deinoxococcus
29	470.5	9.3	713	1	Q05152	Q05152 sulfolobus
30	465.5	9.2	818	10	Q41742	Q41742 zea mays (m
31	464.5	9.2	715	2	Q9K103	Q9K103 streptomyce
32	464	9.2	782	2	Q9L1E9	Q9L1E9 streptomyce
33	456.5	9.0	715	2	Q9X947	Q9X947 streptomyce
34	452.5	8.9	746	2	P72691	P72691 synechocyst
35	413.5	8.2	569	10	Q9XFG6	Q9XFG6 hordeum vul
36	385.5	7.6	656	2	Q9KX51	Q9KX51 vibrio chol
37	379.5	7.5	552	2	Q44528	Q44528 anabaena va
38	322	6.4	332	10	Q9XFG8	Q9XFG8 solanum tub
39	320.5	6.3	600	2	Q9RX51	Q9RX51 arabidopsis
40	295.5	5.8	327	10	Q9XFG7	Q9XFG7 triticum ae
41	295.5	5.8	327	10	Q9XFG7	Q9XFG7 triticum ae
42	284	5.6	559	1	Q55088	Q55088 sulfolobus
43	274	5.4	2032	2	Q45643	Q45643 bacillus sp
44	272	5.4	561	1	P95867	P95867 sulfolobus
45	265	5.2	726	1	Q9P9A0	Q9P9A0 thermococcus

#### ALIGNMENTS

RESULT	ID	Q9XDB5	PRELIMINARY:	PRT:	849 AA.
Q9XDB5	01-NOV-1999	(TRENBLREL. 12, Created)			
AC	01-NOV-1999	(TRENBLREL. 12, Last sequence update)			
DT	01-MAY-2000	(TRENBLREL. 13, Last annotation update)			
DE	PULULANASE TYPE I.				
GN	PULA.				
OS	Fervidobacterium pennivorans.				
OC	Bacteria; Thermotogales; Fervidobacterium.				
OX	NCBI_TaxID=93466;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=DSM 9078;				
RX	MEDLINE=99240435; PubMed=10224005;				
RA	Bertoldo C., Duffner F., Jorgensen P.L., Anttrankian G.;				
RT	*Pululanase type I from fervidobacterium pennivorans vens5: cloning,				
RT	sequencing, and expression of the gene and biochemical				
RT	characterization of the recombinant enzyme.*;				
RL	Appl. Environ. Microbiol. 65:2084-2091(1999).				
DR	EMBL: AF096862; AAD30387.1;				
DR	INTERPRO: IPR00408;				
DR	PROSITE: PS00626; RCO1_2; UNKNOWN_1.				
SQ	SEQUENCE 849 AA; 96591 MW; 1CB6B2A82A351A2 CRC64;				
Query Match					
Best Local Similarity 34.9%; Pred. No. 26-78;					
Matches 339; Conservative 146; Mismatches 318; Indels 167; Gaps 28;					
QY	10	LSGVLVTFWAFNVKQGSAAHNDGNTTIIYHFCPADYQPMSLMMW----	PKDGGAGAEYDF	66	
DB	13	LSVITLITVALST-----SFAE-----TELIIHYHMDGNTGCMIMIMWVEPISREGAAYOF		65	
QY	67	NOPADFGAVASADIPGNSQVGIIVRTQDW-TRDVSADRYTIDLSKGN-EWMLVEGNSQI		124	
DB	66	TE-KIDFGVAAKVPETLTQVGIIVRLGEMREKDVAMDRFIIKDKGAEVMLDGIEDI		124	
QY	125	FNVEKAEADAAKPAVSNALDASNOVLKLSQPLTLGEGXSGTAVDTDTANKDIPVTSVK		184	
DB	125	Y-----TTKP-----DTPSPVL-----		136	

QY 185 DASLGDVAVLAGTFOHIFGSGDMADPNHSTLLKVTNNLYOFGSGLPEGNTQYKVALN 244  
 Db 137 -----FAAARDQYTI--- 146  
 QY 245 DSMNNSYPSDNINLTVPAGCAHYFYSIPSTHAYDTIINNPNADLVQESGVKDLVYVTL 304  
 Db 147 ---EALYTGQVDTT--KVAQAKTVVDCQPLKIAVEKANTDLS-----RTNHAKVVL 193  
 QY 305 GED---PDVSIHTLSIQDTQVQAKQVIRPNVLSNQYYSGGDDLGNTYTKATTEKVAAPT 361  
 Db 194 AEPILKEDVAKMDVQVELEGKPARVIMMELD--KITYDG-PIQGFETPEKTTIRWASPV 250  
 QY 362 STQVNVLLYDS-ATGSYTKIVPMTASGHVWEATVNONLNNYMYMEVYTGCGSTRVAVP 420  
 Db 251 SKTYDVLLYKQMDKEPTQVAPMKYIGNGAMEAVLEGNNMGWFKIKYFSGEYRESDY 310  
 QY 421 YATAIANGTROMIVDLAKTDPAGWNSDKHITPKNT--EDEVIYEMOVROPSIDPNCGM 478  
 Db 311 FSKAVTNKSAKSALIDSKTNPESW--EKVAPPLVAPEDALITYELIHADMTGLDNGCVK 368  
 QY 479 NKGYLALTEKGTGKPDNVKGTGIDSLKQGITVYQALPVPFASNSVDETFDOD---NMGY 535  
 Db 369 NKALYLOLTFEKGTCFPGCVTTGDLHLVELGVTHILPMDFWMTGDEADKDFERSYMWGY 428  
 QY 536 DPNRYDVPECOYATN-ANGNARIKEFEKEMVLSLHREHIGVNMVYVHTFATQI-SDPK 593  
 Db 429 DPLYFTVPEGRYSTDPINPYTRILEVQOMYKALHENGIRVILDMVFPHTGVGVMSPEQ 488  
 QY 594 IVPYXYRTMIOVLIPTDQVLEMK-----LXERPMVQKFIIDSLAKYMWNEYHIDGRFD 648  
 Db 489 AVPYEYFR-----IDKTCATLVNESGCCNVIASERPMMKRYIVDTLAKMWTEYKIDGRFD 543  
 QY 649 LMALGKDTNSKASSELHAINPGIALYGEPMGTGTSALPDQDLTKGAKGMGAVAVENDN 708  
 Db 544 OMGLMDKVTMLAIKSELISKLEPSVVLGEPMGCGAPI---RFGKADVGSTGIAAARND 599  
 QY 709 LRNALDGNVDESSAQGFATGTCITLDAIKNGVESS-----INDFTSSPGETINVTYSHD 762  
 Db 600 FRDLRBSVFNATYKGLMGLAKETGVKRGVAGSIEYDEVIRSFADPQETIYVYVHD 659  
 QY 763 NYTLMDKATLSNPDSSEA---DRIK-MDELAQAVVMTSGVPPMOGCEMLRKKGNDN 817  
 Db 660 NHTLMDKNYLAQAQDITVYKKTTEMLKDAQKLAGAILTISOCIPPLAHAGQDAPARKKFDEN 719  
 QY 818 SYNAGDAVNEFDMSRKAQYPDVFNYSGLIHLRLDHPAFRTTANETINSHLOFLNSPENT 877  
 Db 720 SYKSPISINGLDYARKAEFLIDVFNYKGLIELRKSHIAFRQRTMEDIRKLTFLPSPRKM 779  
 QY 878 VAYELTQVHAKDKKGNITIVYVNPVKTYATINLPSGKMAINATSGKVESTLGOAGESVQV 937  
 Db 780 VAFVLAKD--EKDPKKEILVIYNGDTKDOQFTLPDGTWNVVVDKMDAGTKVLYOVSGKINI 837  
 QY 938 PGISMILHQ 947  
 Db 838 KAISAMVYK 847

RESULT 2

P71095  
 ID P71095 PRELIMINARY: PRT: 668 AA.  
 AC P71095;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE PULULANASE.  
 GN PUL.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteroidia: CFB group: Bacteroidaceae: Bacteroides.  
 OX NCBI\_TaxID=818;  
 RN (1)  
 RP \*SEQUENCE FROM N.A.  
 RC STRAIN=5482;

RX MEDLINE=97113539; PubMed=8955399;  
 RA D'Elia J.N., Salyers A.A.;  
 RT "Contribution of a neopullulanase, a pullulanase, and an alpha-  
 RT glucosidase to growth of Bacteroides thetaiotaomicron on starch.";  
 RL J. Bacteriol. 178:7173-7179(1996).  
 DR EMBL: U67061; AAC44685.1; -;  
 DR INTERPRO: IPR000461; -;  
 DR PFM: PFM00128; alpha-amyase; 1.  
 SQ SEQUENCE 668 AA: 74603 MW: 044D3D62578C6845 CRC64;

Query Match 24.7%; Score 1252; DB 2; Length 668;  
 Best Local Similarity 42.4%; Pred. No. 7e-68;  
 Matches 266; Conservative 100; Mismatches 229; Indels 32; Gaps 10;

QY 348 YTKATTEKVAAPTSTQVNVLLYDSATGS-VTKIVPTASGHVWEATVNONLNNYMY 406  
 Db 47 YTPBATFTLMSPTADVRMLLYEAGGCAHAYETVKMQSEEGTWAIVSKDLIGFTYF 106  
 QY 407 EV---TGCGSTRVAVP--YATAIAPNGTROMIVDLAKTDPAGWNSDKHITPKNI 460  
 Db 107 NVKIDKMQQDT----PGINARAVGVNGKRAAIIIDQSTNPDGMESTRPPLKSPADMI 161  
 QY 461 IYEMDVPEPSIDPNRSGMKNKGYLALTEKGTGPDNVKGTGIDSLKQGITVYQALPVPFAS 520  
 Db 162 IYEMHHRDFSVDTSGKNNKGYLALTEHGTMNSDKLTGIDHLELGVTHVHLPSFDY 221  
 QY 521 NSYDETDPTOD--NMGYDPNNYDVPEGOYATNANGNA-RKEFEKEMVLSLHREHIGVNM 577  
 Db 222 ASYDETFRLNNSYMWGCDPQNNVNPDSYATDQYQATRYKEFKOMYQALHAKGIRVIMD 281  
 QY 578 VYVNHFFATQISDPDKIVPEYVYRTMIOVLIPTDQVLEMKLAEERPMVQKFIIDSLAKYW 637  
 Db 282 VYVNHFFATQISDPDKIVPEYVYRTMIOVLIPTDQVLEMKLAEERPMVQKFIIDSLAKYW 637  
 QY 638 NEYHIDGFRDLMAIGKDTMSKASSELHAINPGIALYGEPMGTGTSALPDQDLTFRG-A 696  
 Db 342 KETHVDFRDLKMGIDHIEITMNEIRKAVNAVDPILCYGMAAEAPQYRADSIAKNGNI 401  
 QY 697 QKGMGVAVFNDNLNMAIDGNVFPDSQAQGFATGTCITLDAIKNGVESS----- 744  
 Db 402 AQIPGVAVFDELDLGDLCGVGOKRKAGFLAGIPGEMSVKFGIAGIEHPVOGDSVNY 461  
 QY 745 --NDFTSSPGETINVTYSHDNYTLMDKIALSNPNDSADAIKMDLAQAVVMTSGVPEM 802  
 Db 462 TOKFWAKQPVQMTSYSCCHGLCLVDRLKASMPDITPEQDRIKDLKLAQAVVFTSGIFPI 521  
 QY 803 QCGEEMLRXKGGNDNSYNAGDAVNEFDMSRKAQYPDVFNYSGLIHLRLDHPAFRTTAN 862  
 Db 522 YAGEIHRDKQGVNDSKSPDAVAIADWRKRTTSADVFMYTKRLIDLRKSHPAFRMGDAG 581  
 QY 863 EINSHLOFLN-SPENTVAYELTDHVNKDKGNITIVYVNPVKTYATINLPSGKMAINATSG 921  
 Db 582 QVRKHLEFLPEGSNLLAFRLKDHANGDHWEDIIAVANSRPTPARLITPVGKTYVVC 641  
 QY 922 KVGESTLGOAGC-SVOVPGISMILHQ 947  
 Db 642 VIDVRLGIONGPEVILIPQSALILYK 668

RESULT 3

Q59319  
 ID Q59319 PRELIMINARY: PRT: 825 AA.  
 AC Q59319;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE ALPHA-DEXTRIN 6-GLUCANOHYDROLASE (EC 3.2.1.41)  
 DE (ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE) (PULULANASE)  
 DE (PULULAN 6-GLUCANOHYDROLASE) (LIMIT DEXTRINASE) (DEBRANCHING ENZYME)  
 DE (AMYLOPECTIN 6-GLUCANOHYDROLASE).  
 GN CALD.  
 OS Caldocellum saccharolyticum (Caldicellulosivor saccharolyticus).





Db 672 TATLALPDEREMAVVCDGRCGCTPPGQARGMLRLDGIATVWL 713

RESULT 5  
ID 09K705 PRELIMINARY; PRT: 717 AA.  
AC 09K705;  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
DE PULLULANASE.  
GN AMYX.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/staphylococcus group; Bacillus.  
OX NCBI\_TaxID=8665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RL Takami H., Nakasone K., Takaki Y.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP001518; BAB06983.1; -; 2A979DFA25638920 CRC64;  
SQ SEQUENCE 717 AA: 81865 MW: 2A979DFA25638920 CRC64;

Query Match 21.4%; Score 1084.5; DB 2; Length 717;  
Best Local Similarity 37.5%; Pred. No. 1,1e-57;  
Matches 232; Conservative 118; Mismatches 255; Indels 13; Gaps 11;

336 QYVYSGDLCNTYQAKTTRKVAAPTSTOVNVLVDSATSGVTKIIVMTASGHCWMEATV 395  
104 QYVYNG-ELGALYSHERTVFAVMPVATEVSLLEYEMHITLAEIYQMKRSPNCWWSVL 162  
396 NONLENNYMYEVGSGSTRAVDPAVTAIAPNCTRGMIYDLTDPAGWNSDKHITPKN 455  
163 DGDLMHCYLRKAVKNGKMTTPVPAKSYVNGEKVIEDLKRTDPAWPTVPIRP-- 220  
456 IEDVEIYEMDVDSIDPNSGKMKRYLALTEKTRGPDNWKGTGIDSLKQGITTHVQLM 515  
221 -EHTIYELHIDATVYPGSHIKKGTYSLSAEONTGSPWCESTGLAYIRELGTHEL 279  
516 PVFASNSVDLPTQD-NMCGDPRNDVPRGQVATNA-NGNARKERKEWLSLAREHIG 573  
280 PVADYGSVDENRREKYNMGDVVYFAPEGSYASDPYDGTREKELKQLAOCORMDLR 339  
574 VMDVYNIHFATQISDFKIVPEYRTMIQVITPDQVLEMLXAERPWKFIIDSL 633  
340 VILDVFNHYVMKESPEKIVPGYFRYDOEGLADOTGCGNDIASRLMKRFIYDCV 399  
634 KYVWNEYHIDGFRDLMAALLGKDTMSKAASLIAINFGIALYGEPTWGTGTSALPDQLT 693  
400 TYMASEYRVDFRDLGIIQDMTKREVKRALAPINRQIILLGEGMDLET-PLENDQKTT 458  
694 KGAQKAGC-VAVFNDNLNMLDGNVPDSSAOGFATGATGLTDAKNGVEGGINPTSSPG 752  
459 LAQRAVPELFFHDFRDAVKGSTFLNKGFCAGNIELKETILASLGEHQAQF-DRPS 517  
753 ETINYVTHSHDYTLMDKIALSNPNDSSEADRIKMDLAQAVMTSGQVPMOGEGEMLRXK 812  
518 QAIQVEVHNDHTLMDKIQVFPDEBEVLRLACRLATVILVTSQGIPLFGGCGEFRRTK 577  
813 GGNDSYNAAGDVAENFDMRRKQAPDVFNYSGLIHRLDPAFRMTTANESINSHLOFLN 872  
578 HGEENSYNSPDMINQDMSRAIVAADEVYKGLIALRAKRAFCSTFSFVNTHMCPK 637  
873 SPENTVAVELTDHYNK-DKMCNIIIVYNPKTVATINLP-SGKMAINATSKVEESTLQO 930  
638 TPSEVYGFHY-HIIAKGPMDEMVLVHNGGDKVDIRLPASGEWEVIVDCKKAGLVPLYT 696  
931 AEG-SVOVPGISMMILHQ 947  
697 FNGHTVSVQSRSSMVLTR 714

RESULT 6  
ID 034587 PRELIMINARY; PRT: 718 AA.  
AC 034587;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)  
DE AMYX PROTEIN.  
GN AMYX.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lapidus A., Galleron N., Sorokin A., Ehrlich D.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RC STRAIN-168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borliss R., Bourisier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
RA Dancloot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Ertian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Film C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
RA Glisz S.Y., Glaser P., Goffeau A., Goldightly E.J., Grand G.,  
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Kleber-Blanchard M., Klein C.,  
RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kunano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Meuel C., Medigue C.,  
RA Medina N., Melledo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Potwollik S., Prescott A.M.,  
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Solio B.,  
RA Sorokin A., Taccori E., Takagi T., Takahashi H., Takemaru K.,  
RA Tachenich M., Tamakoshi A., Tanaka T., Terstila P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,  
RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzengger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
subtilis."  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP008220; AAC00283.1; -;  
DR EMBL: Z99118; CAB1971.1; -;  
DR INTERPRO: IPR000461; -;  
DR PFM: PFM0128; alpha-amylase; 1.  
SQ SEQUENCE 718 AA: 81104 MW: 96F9EB7667877E93 CRC64;

Query Match 20.5%; Score 1039.5; DB 2; Length 718;  
Best Local Similarity 34.3%; Pred. No. 5.8e-55;  
Matches 248; Conservative 125; Mismatches 272; Indels 79; Gaps 19;

234 EGNQYKVALNDSWNNSPDNINILVYPAGAHYTFSTPHTAHVYDTINPNADLVES 293  
54 EAKYK-VCVSD-----HP-----VTFGKI--HCV-----RASS 79

```

QY 294 GVKTDLVTVLGEDPDVSHLSIQDGOVQANQVIRPNVLNSSQYVYSSDDLGNITVYQAT 353
      |||||
      80 GHKTDL-----QIGAVI--RTAAFDDEFYDQ--ELGAAVYADTADT 115
QY 354 TFKVAAPTSTOVNVLVYSATGSVTKIYPRMPSAGHGWEATYVQNLNEMMYVTEYTGGS 413
      |||||
      116 VFKVAAPTATSAAYKL--SHPRKSGRTQMTFLKGVAVYVTDGLHGEYLFECICNNSE 173
QY 414 TRTAVDPATATAPNGTGMIVDLAKTDPAGNSDKHITP--KNIDEVIEYEMDVDESID 472
      |||||
      174 WKEVYDQAKAVTVNGEGVVL---RPDQMKTPA--LKPSHPVDAVYETLHLDKDFSH 228
QY 473 PMSGKNNKGYLALTEKGTGPDNVTGIDSKQDGIHVQMLPVPASNSVDEDTPTOD-- 531
      |||||
      229 ENSGMINKGYLALTEPTDQFANGSSGLAYVKEIGVTHVELLPVNDPAGVDEEKPFLDAY 288
QY 532 NMGYPRVYDPEGOYATNANG--NARIKEFKMVLSTIREHNGVMMDVYVNTFATQISD 590
      |||||
      289 NMGYPRVLFEPFEGSYASNPDPQTRKTELKOMINTLHOGHRLVLDVFNHNYKRENSP 348
QY 591 FDKIPEVYRTMIOVITPTDOV--LEMKLXARPMVOKFIIDSLKYVWVEYHIDGFRPD 649
      |||||
      349 FEKTYPRGIFFR--HDECGRPSNGTGVNDIASERRAKFTIADCYVYWELEYVNDGFRPD 407
QY 650 MALLGKDTMSKASELHAINGIALYGEPMVTGTSALPDQDILLTKGAOKGMVAVFNDML 709
      |||||
      408 LGIILIDIVLVYKKEKATKAKPGIILFEGEGMDLATPLPHQKAAALANAPRMPGIGFENDMF 467
QY 710 RNALDGNVFDSSAGFATGATGTLQAIKNGVGS-----INDFTSPGRTINVTSHDNV 764
      |||||
      468 RDAVNGTFHFKATGFAIANGESAQVAMHIGIAGSSGKALAIPEPSSINVTYESHDNH 527
QY 765 TLMQDITLSPNDSPADRIKMDLEAAYVMTSOGVPEMGGEMLRXKCGNNSYNAGDA 824
      |||||
      528 TFMQDMSALPQENDSKRSKQRLAVAILLQGVPEFINSGGEFRTKQGVENSIOSSS 587
QY 825 VNEFMSKRAQYDVENYVSGLIHLRLDPAPRMTTAMEINSHLOFLNSPENTVAYELTD 884
      |||||
      588 INQIUMDRREFKEDVHYIRLIRSLRKAPAPRFLRSADIQHHECLTLKEHLIAVRLVD 647
QY 885 HYNKRNKGNITVYVNPKNVATINPSC--KVAINATSKVSESTLQAGDEGSOVPGISM 942
      |||||
      648 LDEYEMQDIIYIHASPDSEMRPLNDIPYRLCDPSGEOBDEPT--ELIKTVAVANGIGT 705
QY 943 MLH 946
      |||
      706 VILY 709

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
  Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL, AL355752; CAB90874.1;
SQ SEQUENCE 1798 AA; 194106 MW; 43D13868FACF5C6A CRC64;

Query Match 17.6%; Score 890.5; DB 2; Length 1798;
Best Local Similarity 26.1%; Pred. No. 2.5e-45;
Matches 299; Conservative 149; Mismatches 424; Indels 275; Gaps 40;

QY 24 GOSAHADGNTT-----IIVYFCAGDGYQWMSLMMP--KDCGGAEND 65
      |||||
      702 GRTASATASTGTTPPAEEVPTAASRDYAVVHKRRADGMYDWMGLAWGDLADGENTTWP 761
QY 66 FNPQ--ADSPGAVASADIPGNPSQYIVRTQDMTKDVSADRYIDLSKGENEVLVEGNS 122
      |||||
      762 GTHPTGRATYGAFAVYKLRPGASTYGLVIDKQKQVDAADRTIDYETGEVMIQGE 821
QY 123 QIFYN--EKDAEDAKPAVS-----NAYLDSNOV-LVKSQ----- 156
      |||||
      822 QVTERPEYPAQDTTKAVLHYKRRADGNYDGMGLHWVGDAANPTDMAKPLQPYRTDYGAV 881
QY 157 ---PLTIGSGXSGFYVHDITANKDIPVTSVKDASLGQDVYAVLACFPHI----- 203
      |||||
      882 FEVPLTDGASSLSYIVHKD--EKDLEPTQALDLKADGHEVWLLSGEDHLLPQAGSAAA 940
QY 204 -----FCGSDMAPDNHSTLLKRYNNLKYPSGDPREGNYQY----KVA 242
      |||||
      941 VDLTTSKAWIDRNRYAMNGSDGAS--TQLLASRGTSTABEGRITGDQOYLRLTKSA 998
QY 243 LNDSMNNSYPSDNILNTPAGAHVTSYIPSTHAVYDTINPNADQVYESGVKTDLVTV 302
      |||||
      999 LNDQAKAKPP-----HLKSTAMSV-----DPRDRHREALRGVVA 1037
QY 303 TLGEDPDVSHLSIQDGOVQANQVIRPNVLNSSQYVYSSDDLGNITVYQATFKPKVAPNS 362
      |||||
      1038 QRTATGALLAATGVOTAGALDLIDYADR-----ATRADLAPVRHGRPTLSVAAPTA 1088
QY 363 TQVNVLLYDSATGSVTKIYPMFA--SGHGWEATVQNLNEMMYVTEV-----TQGGSTR 415
      |||||
      1089 QVQKLEIGD-----RTYRMARDATGYWSVYGRPSWKAKARVAVKWPAPRAGEVYTN 1141
QY 416 TAVDPYATATAPNGTGMIVDL--AKTRPAGMNSDKHITPKNIEDEVIYEMDVRFESI-D 472
      |||||
      1142 KVTDPYSALTLTDSERSLVVLDLDRSLAPRGWHEYKRPVAVPLQDAQIDELHIDFSVAD 1201
QY 473 PMSGKNNKGYLALTEKGTGPDNVTGIDSKQDGIHVQMLPVPASNSV-----DETD 527
      |||||
      1202 RTADADRGTYLAFTDKKSDGSRHLR---ELAEAGTSYVHLLPADIATIPERKSEQTT 1257
QY 528 PTPQD-----NMGYPRVYDPEGOYATNANGNARIKEFKEM 563
      |||||
      1258 PQCDLALALPAGSEKQDCVTKAAKAFNMNGYPRVYDPEGSAATDPDQTEVTEFRKM 1317
QY 564 VLSIREHNGVNMDDVYVNTFAT--QISDFKIVPEYVYRTMIOVITPTDQVLEMKLXA 620
      |||||
      1318 VKALINSDGLRVYMDVYVYNTAAAGQAKTSVLDRIVGGYORLLADGSVANSTCCS--NTAT 1376
QY 621 ERMVOKFIIDSLKYVWVEYHIDGFRFDLALLGKD---TMSAASELAINNG-----I 672
      |||||
      1377 ENAMMGKLVVDVYVTAKEKGVDFRFDLGMHPKANIILAVRALDELTLRKGVGDKKI 1436
QY 673 ALYGEPMVTGTSALPDQDILLTKGAOK---CMGYAVFNDNLRLNL--DGNVPS--SAQGA 726
      |||||
      1437 VLYGEGMNGEVA--DARVQVATQKNMAGTGIATSSDRARADVAGGCFPDEDPGVQGRA 1494
QY 727 TG-----ATG-----LTDAIKNGVEGSIINDT----- 748

```

```
Db 1495 SGLIYTPNSSTANGTEAEQKARLLIYQDLIKVGLTNCNLADYSETDTHGKEVKGSEVDYNG 1554
OY 749 -----SSPGETINNYTSHDNTYTLMDKILALSNPNDESA--DRIMDELAQAVYMTSGQVPE 801
Db 1555 APAGYAAAFODALAYADAHNDESLFDALIFPKDTISADBRAMQVLAAMATATLSOGPAL 1614
OY 802 MOGSEMLRKKKGNDNSYNAGDAVNEFDWS-----RKAQYPDVFNV-----842
Db 1615 SQAGTDLRSKSLDRNSYDSGDWFMALHMNCADNGFGRLPPAADNQDKMPYAKPLLGA 1674
OY 843 -----YSGLIHLRLDHPAFKMTTANEINSHLOPLNSPENTYAYELTTHVVK 888
Db 1675 VEYGCPOITGASAAAYRDLRLIRNGEODFSLATYARVQSRLSFPDSEDE-----TPGVIT 1729
OY 889 DKMGNIIVYNN--PKNTVATINLPSGK-----MAINTSGK--VGESTLGOAGSGVOYPGI 940
Db 1730 MRLGDLVYVFNATPERQORVCAAGCTGYRLHRYQAAGADAVYKESAYAAKTOTFTYVPAR 1789
OY 941 SMNLIHQ 947
Db 1790 TVAVFOR 1796
```

```
RESULT 8
OY 09K6N1 PRELIMINARY; PRT; 1072 AA.
AC 09K6N1:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ALKALINE AMYLLOPULULANASE.
GN BH3697.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB001519; BAB07416.1;
SQ SEQUENCE 1072 AA; 120674 MW; 5264235CF24619D9 CRC64;
```

Query Match 16.2% Score 819; DB 2; Length 1072;  
Best Local Similarity 25.6% Pred. No. 2.4e-41;  
Matches 296; Conservative 165; Mismatches 402; Indels 294; Gaps 51;

```
OY 2 AKRLIYVCLSVCLVLTWAFNVKG-QSADADGNTTIIIVHFCSPAGDYQPMSLMMK-----55
Db 6 ASKYSVLEMLVLLVLSFTIMPHNEANADTEADTLRIHORSQDQFENGLMLMGDVRX 65
OY 56 PKD---GGGAEYDFENQPADSFGAVASADIFGNSQVCG-IIVKQDVKDYSADRYIDL--109
Db 66 PSENWPSGGTPEPEKXTD-YGAVYVDIELKEMENIGFVLNATYGDCK-GNCKHVOQLS 123
OY 110 SKGNEVVLVNGNSOIF-----YNEKDA-----EDAKPAVS-----140
Db 124 PEKKEIWIKSGSDVFLRYEVDLRLDHTVRLIHERSDSYHHMKGLMTKKDVASPEHHPSG 163
OY 141 -----NAVYDA-----SNOYLVKLSOPLTLGEGXSGF-TVHDDTANKDIPVTS 182
Db 184 ANFSPDEQTSRYGAYLDVELADANEI-----GFLVYVNRGDKD-----223
OY 183 VKDASLGGDVT-AVLACTPQHIFGSGMDAPDNHSTLKK-----YNNNLXQPSGDLPEG 235
Db 224 -----GDDTFTFDLA-----NHQLEIRGGDDTVYVNPYYQTIEGLVSG 262
OY 236 NYQYKVALNDS-----N--NNSYPSDNIINLTVPAGGAVTESYIPSTHAUVDTINPN 286
Db 263 EL---ISLTREMLRFTSTIWLSEDDLABEQIHLT-DTNGASAVARPHV-----TIH---307
```

```
OY 287 ADLQVESGVKTDLVYTLGEDPDVSHTLSTQTDGYQAKVIPRN--VLNSSOYYVSGDDL 344
Db 308 -----DEKTYILGSPDAE--LAPYHTFADKTYVARAGRRLIDALCYGDGEIG 354
OY 345 GNTYTKATTFKMAWPTSPQVNVLLY--DSATGSYTKYIPMTASGCVMEATYN-----396
Db 355 AVLEHEDGSSTLFEMSPDLADHVSVALDKDOYRIIDDDIMTFGDGQVWQVOTLASENTGL 414
OY 397 QNLENNYMEYVGGSTRTAVDPYATATAPNCTRG--MIVDLAKTDPAGMNSDKHITP- 453
Db 415 DNLAGYTHYHKIRDDDTKLGLDIPYAKSMAAMNNESSYSGKALIVDPSSIGELDPAEI 474
OY 454 ---KNIEDEYIEMDRDPSIDPNSGMKNK---GKYALATEKSGKPDNVTGTGIDSLK 505
Db 475 EGEKREDAITIEVHVHRDFTSDP--GIEBELTAQFGFASFVDK-----LDYIE 521
OY 506 QIGITHYQMLPVPASVSDE-----TPPTDQNNCYDPRNVDYVEGQATN-ANG 553
Db 522 DLGVTHIQLLPWSYVWGBELANHERELAYSSSGNNYMGYDPHSYSLSGMSENPEDP 581
OY 554 NARIKEPKAVLSLHREHIGVNMDDVYNNHFEATOISDFKIVEPYRYTMIQYIIPDQV 613
Db 582 ELRIKEPKHLIDELHRRGMKVIIDVYNNHTAAVEI--FEDLMRYTH--PMDADGTERTSF 638
OY 614 LEMKLAERPVOYKFIIDSLKYVWNEYNHIDGFRFDLMALLGKDTMSKASELHAINGIA 673
Db 639 GGGRLGTHHKMSRRIIVDSITYWEEYKYDGFREFDMGDIDDAETIOIAYDRAKELPNIV 698
OY 674 LVGEPM-----TGGTSALPDDQLTKGAQGMGVAVFNDNLNADGNVFNDSAGGFAT 727
Db 699 MIEGMRITYGDEHGDIMPADO---DMMOHTESVGFSEFENELKSGFGSGQPRFIT 755
OY 728 GATGLTDAIKNGVEGSINDFT--SSPGETINNYTSHDNTYTLMDKIALS--NPNDESDRI 783
Db 756 GGAIRNIEQIFDNLTAQRHNTATNPQGVVRYIEAHNDLTHDVIAQSIKRDPEYHQEIH 815
OY 784 KMDELAQAVYMTSGQVFPWQSGEBMLRXK-----GCN-----815
Db 816 KRIRLGNTVYLTAGTAFHLAAGQEFRTKQFRADTGGQARYKSTYMTDENGKRFVYRFT 875
OY 816 DNSYVNGDAVNEFDWSR---KAQYRD---VFNYSGLIHLRLDHPAFRMTTANEINSHLO 869
Db 876 HDSTIDSSDALINKFDMGKATNKERYPHNLRTYTAGLIELRSTNAFTHGLLDVQANVR 935
OY 870 FLNSPE-----NTVAY-----ELTDH-----YNKDKMGN 893
Db 936 LLDIPELSEIDVLYGATSTDMTGTYYVFNVDTKERTLSLDDLLEGVYIVDSDBAG- 994
OY 894 IIVYVNP-----NKTYATINLPSGKMAINATSGKVBESTLGOAGSGVOYPGIS 941
Db 995 IYAVSPSGFDLTPENETTIPLTAIVLIRVSGDFQENPDHGRPKGG-GKAKGRPHNAGOR 1053
OY 942 MMLIH--QEVSPDHGK 956
Db 1054 GPPRHAGEKGRPHASK 1070
```

```
RESULT 9
P70983
ID P70983; PRELIMINARY; PRT; 1938 AA.
AC P70983;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ALKALINE AMYLLOPULULANASE.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=KSM-1376;
```

```

RX MEDLINE-96394539; PubMed-8798645;
RA Hatada Y., Igarashi K., Ozaki K., Ara K., Hitomi J., Kobayashi T.,
RA Kawai S., Matabe T., Ito S.;
RT "Amino acid sequence and molecular structure of an alkaline
RT amylopullulanase from Bacillus that hydrolyzes alpha-1,4 and alpha-1,6
RT linkages in polysaccharides at different active sites.";
RL J. Biol. Chem. 271:24075-24083(1996).
DR EMBL: D78258; BAA11332.1;
DR HSSP: P06278; IYTS.
DR INTERPRO: IPR000461;
DR INTERPRO: IPR000834;
DR INTERPRO: IPR001899;
DR PRAM: PF00128; alpha-amylase; 2.
DR PRAM: PF00746; Gram_pos_anchor; 1.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN.1.
SQ SEQUENCE 1938 AA; 214917 MW; A29D3F715EB6A72E CRC64;

Query Match 16.28; Score 817.5; DB 2; Length 1938;
Best Local Similarity 25.88; Pred. No. 7.4e-41;
Matches 284; Conservative 153; Mismatches 369; Indels 293; Gaps 50;

QY 35 TIIVHYCPADYQPMSLRMV-----PKD---GGGAEYDFNQPADSPGAVASADIRGPS 86
DB 835 TLRHYQFTDNTSYENLGLMLGVAAPSENNPSSGCTPQAGNVTD-YGAYVDVELAEDAQ 893
QY 87 QVG-IIVFTQDMTRKDVASADRYIDLKSG--NEWMLVEGNSQIF----- 125
DB 894 NIGGLVLTNTGDDK-GGDKAVLELSPDLNEIMIKGSDVEFLVPLVPLANTVRIHER 952
QY 126 -----YNEKDAE-----DAAKPAVSNAYLDAASNOVLVKSPLTLGEGX 164
DB 953 TNADYEGGLNMWEDVESPDGMPGADAAIGIKYGAIVD-----IKLK----- 997
QY 165 SGFVHDDTAKKDIPIYSYVKRNASLQ--DYTAVALGTFQHF--GGSMPAPNHSITLKK 220
DB 998 -----EDAK-IGFLFVNKQSGQCTGDMTFMLKQYNLFVKEGD-----K 1038
QY 221 VTNNLYQFSGDLPEGNVQYKVALDNNNSYPSDNINLTV-----PA 262
DB 1039 VYTNPY---GVPLALVSGEV-LSDKL-----ISLFTREGLDEELKEQLEIKDV 1086
QY 263 GGAHVTSYIDSTHAIVDTINNPNADLOVESGVKT-----DT---VTYTLGSDPDVS 311
DB 1087 DGNVSEF-----DVTIE-GEKTVHGEFDELEKIPFSVY-----LD 1123
QY 312 HTLSIQGTQYAKOVIIPNNVLNNSQYSSGDLGNTYOKATTFVMAPTSQVAVLYD 371
DB 1124 RTISVKS-GMK-----LIDEMVAVDGKLGAEIHEDGTATLKVWSPKADNVSVLYD 1173
QY 372 SA-TGSVYRKIPMTASGHGVMEATVNO-----NLNNYMYTEVYGQSTRAVDPYATA 424
DB 1174 KVDQNEVVDTEIMYKGDGWSVKLTIKDNTGDLGSLKGYIHHETHIGDNTNALDIPYAKS 1233
QY 425 IAP-NGTRG-----MIVDLAKTDPAGNNSDKHTP--KNIEDEVIYEMDVDFSDIPNS 475
DB 1234 MAAMNNNEAGDYVGAAYIDIGSICP---ELDYADIPGEKREDTIYIEVHVDFTSDPNI 1290
QY 476 GKNK-----GKYLALTEKCTKGDNNKTGIDSLKQIGITHVQMLPV---FASNSVE----- 525
DB 1291 GBDLAAQGTGTRASFVEK-----LDYIQELGVTHIQLLPVMYSYSENEFSGERM 1339
QY 526 -----TDPTQDMWGDYPRNYDPEGOYATN-ANGNARIKEFEKEMVLSLREHIGVMNDVY 580
DB 1340 LEVASTGTNNMKGYPHNYFSLSGMYSSENPEDPELRKEFKLLINEIHRKMGVLDVVF 1399
QY 581 NHTFATQISDQKIVPEYYVFTMIOVIITPDQVLEMKLAERPMAQKFIISLKTAVNEY 640
DB 1400 NHT--AQVHIEFDLVPNYH-FMDADGTPRTSFGGGRGLGTTHEMSRRLVLYDSIKIHVDEY 1456
QY 641 HVDGFRPLMLGKDTMSKASELHAINPGIALGEPPT--GGTSALPDDDLTKGAK 698
DB 1457 KVDGFRPLMGDHDADESIOLAPDEAKKLIPNIVMIGEGVYTPAGDGEFVQAADQOMQY 1516

```

```

QY 699 GMGVAVFNDNLNMGVNFVDSAGFATGATGLTDAIKNGEVSINDF-TSSPETINY 757
DB 1517 TEAVGSFDEFENELKSGFSGCGPRFITGCAVNVQOIFDNIKADPHNMADQPDVQY 1576
QY 758 VTSHDNTYLMKDIALSNPDSRADRIKMD-----ELAAVNVTSQGVPPMGCEMLRKX 812
DB 1577 IEAHNDNLTYLVIAOSIKKDPLEAENDEIHKRIYVGNAMVYLSQGTAFLAHGOEFGTRK 1636
QY 813 -----GKN-----DQSYNAGAVNVFEDMSRRK---QYDVEN- 841
DB 1637 QMRAPATAPKSTYMTDADGNPFYPIFTINDSYSSDILNFMFKATDAKRTY-VNNV 1695
QY 842 ---YSGLIHLRLDHPARMTTANBINSHLOFNSPE-----NTVAYELTDH 885
DB 1696 TRDYTAGLIELLRSSDAFLGRELVDNSVMTVADAEIKEDQLVVAYSRSVTAQVEYYTF 1755
QY 886 VN-----KDKGNIIYVNPNTVATINLPSG-----KKAINATSGKGE 925
DB 1756 VNADYSSRTLGLDGLTEGVVVVDAEAVNAGVAPAGFELTAEGITLPLTVVVRGE 1815
QY 926 ST-----LGOAGSVQVPE 939
DB 1816 QEGTDPGDDGDNTPPG 1834

```

## RESULT 10

```

ID Q9RXB0 PRELIMINARY; PRT; 910 AA.
AC Q9RXB0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ALPHA-DEXTRAN ENDO-1,6-ALPHA-GLUCOSIDASE.
GN DR0405.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI.
RX MEDLINE-20036896; PubMed-10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodsan R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathavan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
RT radiodurans RI.";
RL Science 286:1571-1577(1999).
DR EMBL: AE001900; AAF09983.1; -.
DR TIGR: DR0405; -.
SQ SEQUENCE 910 AA; 97926 MW; FEAA46607BF9A504 CRC64;

```

## Query Match

16.18; Score 815.5; DB 2; Length 910;  
Best Local Similarity 26.98; Pred. No. 3.1e-41;  
Matches 260; Conservative 126; Mismatches 332; Indels 247; Gaps 34;

```

QY 87 QVGIIVTQDMTRKDVASADRYIDLKSGNEVWLVEGNSQIFYENKDAEADAAKPAVSNAVLYDA 146
DB 7 KLGFTYHKGD-DKDPEDGTLTREGQNLMIYSGNTPSTRTDR-----VRTVGDL 58
QY 147 SNQVLKLSQPL-----TLCEKXSGFTVHDDTANKDIPYTSVADASLGQDYTAVALAGTQ 201
DB 59 KQOQALMSRDLILVYKPAALQPGAWLTLH-----VSPATLKLIGAD----- 99
QY 202 HIFGSDMAPDNHSTLLKRYTNNTLYQFSGDLPEGNVQYKVALDNNNSYPSDNINLTVP 261
DB 100 GVSQGLTLA-----LEEVEGGL-----TP 118

```

QY 262 AGAAHTFSTYIPSTHAYDTINNPNADLOVESGKTDLTVTTLGDEPDVSHLS-----IQ 317  
DB 119 ALQAKAPYL-----AGYLL-----LQIRADRAKLPTALSGO--LAVSSVLPGGKVLID 164  
QY 318 TDCYQAKQVYPRNVLSQSYYYSGDLGNTYTKATPFKMAPTSTQVNVLLYDSATGSV 377  
DB 165 ATGVQVAMAL-----DDLXAYAG--PLGVTWQGNKPTVILMAPTADVY--LHVSRRDGTQ 216  
QY 378 YKIVPTASGHGWEATVNONLENNYMYEV-----TGQSTRTAVDPYATAIAPNGTR 431  
DB 217 ERLPMTBPTGQWTFASGQSMWGLAYREVRFAFSTGNIETNLVTDPYSVALTRNSTH 276  
QY 432 GMIVDLAKD--PAGNSKHIITPKNIEDEVLEYEMDVDPES--IDRPSGKNKKKYLALTE 488  
DB 277 SYPADLSDNSQKPAQNNALKKPALNSFADLSFEYLHLRDFSAADASVPAARCTYLAFTK 336  
QY 489 KCTKGDNVKTCIDSLKQIGITFVQLMPV-----518  
DB 337 PPSDG-----MTHLRALAGLAKAVHLLPFDLTATINEDKGQOTPGDITQFGPNSSEQK 392  
QY 519 ASNSVDETPDQDNMGVDPFRNYDVPQYATNANGNARIKEPEKMYLSLREHIGVNMVY 578  
DB 393 AVNAVADADAY--NMGYDEYHSMVPEGSYAVNP--AERTKEYROMYMSLMAAGLRYVQVY 448  
QY 579 VYNHTEAT---QISDFDKTYPEYVYRTMIOVILPTQVLEMKLXABRPVOKITISLKY 635  
DB 449 VFNHTAASQAKERSYLDKTPGYTHRLNMGVENSTCCS--NTATHTTKMRLMVTPLVY 507  
QY 636 WVENYHIDGFRDLMLALGKDTMSKASELHAINPG-----IALYGPWPGTSAIP 687  
DB 508 MAREYVDFGRDLMGHINWADQMARLALDLTPAKDGVQKQIITLYEGMDFG-----562  
QY 688 DDOULTKGAOK-----GKQVAVFNDNLNAL--DGNVFPDS--SAQCPATG--AT-----730  
DB 563 EYQGNRGNANATOLNLYCGAGIGTFNDRYDALRGSGFGLQEOGATGATVPCNQAGN 622  
QY 731 -----GLTDAIKNGVEGSIND-----FTSSPGETINYT 759  
DB 623 DSAEKKGQIADLYKLGLAGLRDYRTDHSKGFVYTGQGLRYGDPAPGYASAPRETTINVAS 682  
QY 760 SHDNYTLMDKIALSNP--NDSEADRIKMDLAQAVVMTSGVPEMGGEEMLRXKKGNDNS 818  
DB 683 ANDNQDLMADVLLKTPANNMSAARVFNQNLAIISYLLGGGLPVPVAGDELRSKSFDTOS 742  
QY 819 YNAGDAVNEFDS-----RKAQYFDVFNYS 844  
DB 743 YNSGDMFNLDMYGTNGFGKGLPRAEKNBANMALRYTLGDAALVTPADITRASDHFR 802  
QY 845 GILHLRDLPAFRMTTANEINSHLOFLNSPENTVAYELDHYN--KDKMGIIYVYVNPNT 903  
DB 803 ELURVSSSLRRLDTAAGVQOSLSTFLPAPGVYIANKLSGVSATVPYRDLVIFNGSAD 862  
QY 904 VATIN 908  
DB 863 AVTLN 867

RESULT 11  
ID 064454 PRELIMINARY: PRT: 988 AA.  
AC 064454:  
DT 01-AUG-1998 (TREMBLREL 07, Created)  
DT 01-AUG-1998 (TREMBLREL 07, Last sequence update)  
DT 01-MAY-2000 (TREMBLREL 13, Last annotation update)  
DE STARCH DEBRANCHING ENZYME.  
OS Oryza sativa (Rice).  
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:  
OC Magnoliophyta: Liliopsida: Poales: Poaceae: Oryza.  
OX NCBI\_TaxID=4530;  
RM (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, NORIN-8;  
RA Francisco P.B., Zhang Y., Park S., Ogata N., Yamanouchi H.,

RA Nakamura Y.,  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB012915; BAA28632.1; -  
DR MENDEL; 28470; Oryza; 2480; 28470.  
DR INTERPRO; IPR000461; -  
DR PRAM; PR00128; alpha-amylase; 1.  
SQ SEQUENCE 988 AA; 109858 MW; 90537693EF15A4AE CRC64;

Query Match 15.9%; Score 805; DB 10; Length 988;  
Best local similarity 27.6%; Pred. No. 1,5e-40;  
Matches 252; Conservative 137; Mismatches 273; Indels 250; Gaps 37;

QY 246 SWNS-----YPSDNLIVPAGAH-----VTFSYIPSTHA 277  
DB 96 ANVNVQDTSLEFLYASRDAIMHVSIDCAIHGYDSKIELEPEHASLPDVAEKFPFIRS---152  
QY 278 VYOTINNPNADLOVESGKTDLTVTTL-----GEPPDVSHLSIOTDGYQAKQVYPRNVLS 334  
DB 153 -YTFVRVPS--VDVASLVCQLAVASYDAHGRHODVT-----GLQLPGLV-----D 196  
QY 335 SOYYSGDDLGNNTYTKATPFKMAPTSTQVNVLLYDSATGSVTKIVPTASGHGWEAT 394  
DB 197 DMFAATG--PLGAVFSQKDVLYLMAPTAODRYCFTDGRGPILOQYOLKEL--NGWSTV 254  
QY 395 VNONLENNYMYEV-----TGQSTRTAVDPYATAIAPNGTGMIVDL--AKTPAGWN 446  
DB 255 VPRPREQYVLYEVKHYHPTSOVEKCLADDPYARGLSANGCTRWLVDINSETLTKPASMD 314  
QY 447 -----SDKHIITPNIEDEVLEYEMDVDPES-----IDPNSGKNKKGYLL-----466  
DB 315 ELSDERPNELESFSDISLEYELHIDFSHSDTVDCS---RGGFVHLHFFRLNLLNDF 370  
QY 487 -TEKGTGKPDNV-----KTGIDSLKOL---GITVQLMPVPAASNVDET-----526  
DB 371 CSPPTIKHPGRIMETWQDSAGIRHLRLKLSAAGLTIVHLLPSRHPFASVDKNSMKFVDE 430  
QY 527 -----DPTOD-----NMGYDPFRNYDVPQYATNANGNARIKEPEKMYL 565  
DB 431 AOLAKLPQSDQOAAIVSIQOEDPYNMGYDPVLMGPKGSYASNPDCPSRIILEYROMVQ 490  
QY 566 SLREHIGVMDVYVYHNTFAT---QISDFDKTYPEYVYRTMIOVILPTQVLEMKLXAE 621  
DB 491 ALNRIGLRVMDVYVYHNLDSGPGVSSVLDKTPGYTHRLNMGVENSTCCS--NTATHTTKMRLMVTPLVY 507  
QY 622 RPAVOKFIIDSLKYWNEVYHIDGFRDLMLALGKDTMSKAS--ELNAI--NPGIA 673  
DB 550 HFVVDRLTYDDLLNMAIVYVDFGRDLMGHIMKSTMIKAKSAIRSLTRBYHGYGSKIY 609  
QY 674 LYEPWTTGTSA-----LPDDQLLTGAQKQKGVAVFNDNLNALD--GNVFDSS--AOGFA 726  
DB 610 LYEGGDVDFEVAQNKRCINASQI---NMSGTIGSFNDRIIDSVNGCNFPGNPDQGF 665  
QY 727 TG-----ATGLDAIKNGVEGSIND-----746  
DB 666 TGLFLEPNGYGCGNEADTRRELAT--YADHIQIGLAGLRDYLRHTGAEKKKSDITYFD 724  
QY 747 -----FTSSPGETINYTSHDNYTLMDKIALSNPNDSEAD--RIKMDLAQAVVMTSGVY 800  
DB 725 GSPVGYTSSPVEITINVSADNETLFDIYSIKTPIGISIDEKRIHMLASSMIALOGIP 784  
QY 801 FMOGGEMLRXKKGNDNSYNAGDAVNEFDS-----831  
DB 785 FFHAGDEILRSKSLDRDSYNSGDMFNLKFTYETJNNMGVGLPRDRKNEEMHMLIKPRLN 844  
QY 832 -----RKAQYFDVFNYSGLIHLRDLPAFRMTTANEINSHLOFLNSPEN---YAYEL 882  
DB 845 PSFRPLKNHILSAFDNEVDLTKIRYSSPLRLSTASDIEQVRVHNTGMPVGVYVMSI 904  
QY 883 TDHVNK-----DK--WGNITVYV--PNKT-----VATINLPSCGMATNATSGKGS 926  
DB 905 KDAQNEKCKMAQLDKNFSYVYVITFNVCPIHEYSIEIHDLASLGLELPIQVYNSSDALYROS 964

QY 927 TLGOAGSVQVP 938  
DB 965 AYEASGRFTVP 976

## RESULT 12

ID 081638 PRELIMINARY; PRT: 962 AA.

AC 081638;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, last sequence update)  
DT 01-NOV-1998 (Tremblrel. 12, last annotation update)  
DE PULIDANASE-TYPE STARCH DEBRANCHING ENZYME.  
GN ZP01.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-ENDOSPERM;  
RA Beatty M.K., Rahman A., Myers A.M., James M.G.;  
RT "Genetic and biochemical characterization of ZPU1, a pullulanase-type  
starch debranching enzyme from maize."  
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF080567; AAD11599.1;  
DR MENDEL; 32273; Zeama; 2480; 32273.  
SQ SEQUENCE 962 AA; 106239 MW; 5F5AAB7C6EBD5D10 CRC64;

Query Match 15.8%; Score 797; DB 10; Length 962;  
Best Local Similarity 27.0%; Pred. No. 4,4e-40;  
Matches 262; Conservative 143; Mismatches 325; Indels 240; Gaps 38;

QY 134 AAKPAVSNAVIDASNQVLYKSLPILTEGKSGFTVHDTAKNDIPYISVKRASLQDVT 193  
DB 56 APRPMATAEBSAGSDVGV-----AAVESAGFLL-----DARAYVTK 94  
QY 194 AVLAGTFOHIFGSDM-APDNHSTLKLKVTNNLYOFSGDLPENGYOKVALDMSNNSYP 252  
DB 95 SLIA-----WNTSDOKTSLFLVASRNATMCMSSODKKGVDSKVELDPE-NDGLP 142  
QY 253 SDNINLTVPAGAHVT--FSTYPTSHAVYDTINPNADLOVESGVKTDLVTVTL---GED 307  
DB 143 SS-----VTOKFPFIS-----YRAFRISS--VDVATLVKCOLAVASFDAHGR 185  
QY 308 PDVSHSTLSIQTDGQAKQYIPRNVLNNSQYSSGDDLGNTYQAKTFYVMAPTSTQVAV 367  
DB 186 QDVT-----GLQPGVL-----DDMFAYTG--PLGTTFSEEVASMTLMAPTADVASV 230  
QY 368 LLYDSATGSVTRKIVPMTASGHCWMEATVQNLENVYMYEV-----TGCGSTRFVADPY 421  
DB 231 SFYDPAAPPLETVOLEL-NCWMSVGTGRNENRYLYLEVYVYHOTTCINIEKCLAADY 289  
QY 422 ATAIAPNGTQGNIVDL--AKTDPAGWNSDKHITPK--NIEDEVIYEMDRDESI-DPNSG 476  
DB 290 ARGISANSTRFVLVDINNETLKLPLAMDGLAAEKPRIDSPDISIYELHIDRPSAHSTVD 349  
QY 477 MNKNGKYLALTEKGTGKGDNVKTGIDSLKQLDITHVOLMPVANSVD-----524  
DB 350 CPEFGGCAFTFQDSVGEHLK-----KLSDAGLTHVHLLPSFGVGVDIDIKSMKCVDEI 405  
QY 525 -----ETDPTQDMNGDPRNVYDVEGQYATNANGNARIKEKKEW 564  
DB 406 ELISKIPSPSDLOQAALVAIQEDPY--NMGYNPVWNGVYKGYASNPQPSRIEYRLAV 463  
QY 565 LSLAREHIGVNDVYVNTFATQ---ISDFDKIVPEYRYRTIMQVYIIPDQVLEMKLXA 620  
DB 464 QALNRLGLRVVADVYVNHLYSSGPAITSVLDKIVPGYLLRDSNGQTFENSAVAV-NFAS 522  
QY 621 ERPWQKTLISLKYWVNEVHIDGRFDMLALGKDTMSKASELHAI-----NPGI 672  
DB 523 EHFVNDRLIVDOLLMAVAVKYVDFRDLGHIKMTIMRAKSALOSLIDIEHGVGSKI 582

QY 673 ALYGEPMGTGTSALPDDQLITKGAQ---KMGVAVFENDMLRNALDGN--VFDSSAGGFAT 727  
DB 583 YLYGEGWNGEVA--ENQNGINGSQLMNSCTGIGSFNDRIKRAINGSPFGNPLQGGFST 640

QY 728 G-----ATGILT-----DAKNVGECSIND-----746

DB 641 GLFLERNGFYQGNETETRLTATYADHIOIGLAGNLKDDVYVISHTEARKSEIRTFDGS 700

QY 747 ---FTSSPETINIVYVSHNNTLMOKIALSNPNSDAD-RIMDELAQAVVMTSOGVPM 802

DB 701 PVGIASPLETINIVASAHNNETLFDLISLKTPODLSIDERCINHLSTSMIALSOGIPPF 760

QY 803 OCGEEMLRXKGGNDNSYNAGDAVNEFDMS-----RKAQYR-----837

DB 761 HMGDELLRKSILDRSDYSGDMFNKIDFTYETNNMGVGLPPREKNGSMPLKKPRLNPS 820

QY 838 -----DV---FNYSGLIHLRLDHPAPMTTANETNSHLOFLNSPEN---TVAYELTD 884

DB 821 FRPAKHDIYALDKFIDILKIRYSSPLRLLTASDIVQHVHFHNTGPSLVPGIYVMSIED 880

QY 885 HYN-----KDKMGNIIIVYN--PNKT-----VATINLPBGKAINATSKVGESE 928

DB 881 ARNDRDMAQIDETESCVAFTENVCPEYSEIPLDASLRLQLHPVQVNSDALARQSAV 940

QY 929 GOAEGSVQVP 938

DB 941 DPAATGRFTVP 950

RESULT 13  
P93416  
AC P93416; PRELIMINARY; PRT: 986 AA.

DT 01-MAY-1997 (Tremblrel. 03, Created)  
DT 01-MAY-1997 (Tremblrel. 03, last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, last annotation update)  
DE STARCH DEBRANCHING ENZYME PRECURSOR.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-CV, FUJIKIRARI; TISSUE-ENDOSPERM;  
RX MEDLINE=96275956; PubMed=8680309;

RA Nakamura Y., Umemoto T., Ogata N., Kuboki Y., Yano M., Sasaki T.;  
RT "Starch debranching enzyme (R-enzyme or pullulanase) from developing  
rice endosperm: purification, cDNA and chromosomal localization of the  
gene."  
RL Planta 199; 209-218 (1996).

DR EMBL; D50602; Orysa; 2480; 16402.  
DR INTERPRO; IPR000461; -

DR PFAM; PF00128; alpha-amylase; 1.  
FT SIGNAL.

FT CHAIN 1 74 POTENTIAL.  
FT SIGNAL 75 986 STARCH DEBRANCHING ENZYME.  
SQ SEQUENCE 986 AA; 109592 MW; FBA89E134CAD27 CRC64;

Query Match 15.7%; Score 793; DB 10; Length 986;  
Best Local Similarity 27.6%; Pred. No. 8e-40;

Matches 251; Conservative 134; Mismatches 277; Indels 248; Gaps 37;

QY 246 SWNNS-----YPSDNLNLTVPAGAH-----VTFSYIPSTHA 277

DB 96 ANVNVNDQDTSFLFYASRDATMHVSDGALIGYDSKIELEPEHASLPDVAEKEPFIS--- 152

QY 278 VYDTINPNADLOVESGVQVTDLVTVTL---GEPDVSHLSIQTDGQAKQYIPRVNLS 334

DB 153 -YRTERFPSS--VDVASLVKCOLAVASVDAHGRQDVT-----GLQPGVL-----D 196

OY	335	SOYVSGDGLDNITOKATTFKWAAPTSTQVNVLLYDSATGSYVKIYPMPTASGHWGAT	394
Dd	137	DMFATVG-PLGAVPSDKQVDYLIAFPAADQRYVCFYDGPAGPLIQTYQLEL-NGWASYT	254
OY	395	VNQNLENNYVAYEV-----TGGSSTRAADPAVATAIPAGTSGMIVDL--AKTPAGWN	446
Dd	255	VPRVPENOYVYLYEVKVVYHPSTSQVEKCLADDPYARGLSAAGCTRMWLDINSETLKPASMD	314
OY	447	--SOKHITPKNIEDEVIYEMOVAPRES-----IDNSGMKNKXGKVAL-----	486
Dd	315	ELSDPEPLLESFSDSIYELIIRDFSAHSDIVDONS-----RGCVHLHFFRLFRLLNIDF	370
OY	487	-TEKTKCPDNV-----KTGIDSLNQL-----GITHVQLMPVPASNSVDET-----	536
Dd	371	CSPPITKHPGRIMETVMODSAGIRILRLKLSAAGLTHVLLPFIHFASVDDNKSNNKPFVE	430
OY	527	-----DPTOP-----NMGXPRAYVDYDEQOYATNAGNARKIEFEKENVL	555
Dd	431	AOLAKLPLPGSDBOQAAIVSIQOEDPPYMWGVDYVLMGVPKGSVSNPCPFRILIYEROMVO	490
OY	566	SLHREHIGVNMDDVYVNIHTFAT-----QISDEPKIYPEVYRYFTMIOVILIPTQVLEMKLXAE	621
Dd	491	ALNRIGLVAVDDVYVNIHSDSGPRGVSSYLDKITYPGYILRRNVNGQIENSAANN-NTASE	549
OY	622	RPNVQKFIIDSLKTIKTVANEYHIDGFRFDLMALLKGDYTKSAAAS-----ELHAI-NGCIA	673
Dd	550	HEPMYDLRLVDDLLMVAIYKVDGFRFDLMGHIMKSTMIKAKSAIRSLTRPVHGVGYSKIY	609
OY	674	LYGEPWTTGTSA-----LPDQULLTKGAOKMGVAVPNDILRNALD-GNVFDS-AGFA	726
Dd	610	LXGECMDGCEVAOKNRGINASQI-----NMSTGICGSNDRIRDSYVNGCNPFQNPLOOFS	655
OY	727	TG-----ATGLTDAIKNGVEGSIIND-----	746
Dd	666	TGLELEPNQYVQGNEDATRRRELAT-YADHIQIIGLAGMLKVLVLRTHGEAKKGSDIYFD	724
OY	747	-----FTSSPGETINVYVSHDNTLMLKIALSNPNDEAD-RIKMDLAAVYVITSGVP	800
Dd	725	GSPVGTSTSPVETIINYVSAHNETLEFDVLSKTIPIGLSIDECRINHLASSMIALSOGIP	784
OY	801	FMQGEEMLRKKGKGNDSYVNAAGDVAVNEFDMSR-----	832
Dd	785	FFHAGDEILRLKSLSDROSYNGDMFKKLDLHMNPPICGRLLQELIRMKNMHLIKPRLNPS	844
OY	833	----KAQYPPDYFNYSGLIHLRLDHPAFRMTTANEINSHLOFLNSPN-----TYAVELTD	884
Dd	845	FRPLKNHILSCFEDNEFDLKIIRYSPPFLRISTASDIORVAFINGTGSMVPGVIVMSIKD	904
OY	885	HVKN-----DK-MGNITIVYV-N-PNKT-----VATINLPSSGKMAIINATSGVGE STL	928
Dd	905	AQNECKMAQOLDNYSYVYITFENVCPIHEVSTIEIHDLASLGLELPIQVNSSDALYRQSAY	964
OY	929	GOAEGSVOP 938	
Dd	965	EASKRCFTVP 974	
RESULT 14			
O9S7S8			
ID	O9S7S8	PRELIMINARY:	PRT: 962 AA.
AC	O9S7S8:		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-MAY-2000 (Tremblrel. 13, Last annotation update)		
DE	LIMIT DEXTRINASE.		
OS	Hordeum vulgare (Barley).		
OC	Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.		
OX	NCBI_TaxID=4513;		
RA	SEQUENCE FROM N.A.		
RA	*MEDLINE=99169277; PubMed=10069825;		
RA	Burton R.A., Zhang X.Q., Hrnova M., Fincher G.B.;		

RT "A single limit dextrinase gene is expressed both in the developing  
 RT endosperm and in germinated grains of barley.",  
 RL Plant Physiol. 119:85-87(2(1999)).  
 DR EMBL: AF120500; MD3448.1; -;  
 KW EMBL: AF122049; MD3447.1; -;  
 KW Hydrolyase, Glycosylase.  
 SQ SEQUENCE 962 AA; 105752 MW; 5158386FAED162EE CRC64;

Query Match	Score	DB	Length
15.68	787	10	962

Matches 232; Conservative 134; Mismatches 282; Indels 188; Gaps 28;

OY	269	FSYJIPSTAAVYDTJNNPADQVSSGKTDLVYVTLCEDDPVSHTSIJODGVOAKQYIP	328
Dd	149	FPFJSS-----TRARRBSS-VDAVASLKCOLVASFQADGK-----HVDYTGJQLPEVL-	197
OY	329	RNVJNSSQYVYSGDDLQNTYQKATPKVVAPTSTOVNVLLYDSATGSVTKIYPMPTASGH	368
Dd	198	-----DDMFAYTG-PLGAVFSEDSVSLHLMAPFAQGVSCFPDGPAPALETVOLKES-N	250
OY	369	GWMEVTVQNLNENNYMEVYVGGOSTRAV-----DPYATAPNPGTRGMYDL--AKT	440
Dd	251	GWMSVYTGREMENRYLYLEVQVYHPTKAOYLKCLAGPYARGLSANCARMLVDINNETL	310
OY	441	DPAGNSDKHITPK--NIEDEVITEMVODRPSI--DPNSGKMKCKYALATEKGCKPDVY	497
Dd	311	KPASDELADERPKLDSSTDTITIELHROPSAIDGTVDSKRAFAPAFVQASAGHEIL	370
OY	498	KTGJDSLKQLGJITVHQLMPVPAFASVD-----E	525
Dd	371	R-----KLSDALGTJLHVHLLPSFHPAGVDDIKSNMKFVDECELATFPQPSDQMOAAVVAIOE	426
OY	526	TDPPQDNNGYPRNYTDVPEGOYAINNAGNMAIKFKFKEMVLSLHREHJGVAMDVYVYNTFA	565
Dd	427	EDPY--NMGYNPVLWGVKPGYSADPQPSKILERYOMVALNLTIGLRYVADVYVYNNLDS	484
OY	586	TO-----IDPFKIPYEVYVYFTMIOVILPTDOVLBMKJXAEPRYQKFIJLSLKWVNEYH	641
Dd	485	SGPGSIVSLDKIYPGYVYRDTNGOJENSAAHN--NASEHEFVYDRILYVDLLMNAVNYK	543
OY	642	IDGFREDIMALLGDTYMSKASSELHAINPG-----IALYGBPWTCGTSALEPDDQLT	693
Dd	544	VDFGFEFLMGHIMKRTMMRAKSALOSLTTTDAHGVDSKIYLYGGM--DFAEVARNQGI	601
OY	694	KGAQ---KGMVAAPFNDLKRALD--GVTFDS--AQGATG-----	728
Dd	602	NGSOLNMSGTIGISGNDRIDALINGCNPFGNPLDQGFNTGJLEPNGFYQNEADTRSL	661
OY	729	ATGLDIAKNGVEGSIND-----FTSSPGETINYVTSHDN	763
Dd	662	AT-VADQJQJOLAGLRBYVYLISHTEGAKKSEIHTFDGLPYGTATPIETINVSASHDN	720
OY	764	YTLMDKIALSPNDSAD-RIKMDLAQAVVYTSQVYPMQGGEBMLKXKGNDNSTYAG	822
Dd	721	ETLEPVYISVKTPTMLSYDERCIRINHLLASSMMAJLQGIPIFHAGDEILRSKSIDRDSYNG	760
OY	823	DAVNEFDS-----RKAYPVVNYSSLIH	848
Dd	761	DMFNKLDTTYETNNMGVGLPPESEKNEDNMPLMKRLLENPSFKPAKGHIALALDSFYDILK	840
OY	849	LRLDHPARMTTANEINSHLOPLNSPENTVAYELTDHVNKDK-----WGNJIV	896
Dd	841	IRYSPLERLSTANDIKQVRAFHHTGQSLVPGVYVMGIEBARGESPMAOLDTNFSVYT	900
OY	897	VYN--PNK-----TVATINLPSGKMAINATSGVYGBSTLQAGCSQVYPGISMTI	944
Dd	901	VFNVCPRHEVSMIDIPALASMGELPAPVOYNSDTLVRRKSAYESATCRAFTVYGRVTSV	956
RESULT	15		
ID	048541	PRELIMINARY;	PRT; 904 AA.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 08:57:29 ; Search time 70.9 Seconds

(without alignments)  
879.620 Million cell updates/sec

Title:

US-09-262-126c-4

Perfect score:

5633

Sequence: 1 MKYTRNALVIGSLVLLSGC.....AVLELPQGEAQAAGLPVSKR 1091

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0401.\*  
1: /SIDS2/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4773.5	84.7	1096	9	P82507
2	1503	26.7	964	17	R98371
3	1419.5	25.2	878	20	W81973
4	997.5	17.7	558	18	W34878
5	840.5	14.9	862	21	Y78513
6	820	14.6	922	19	W37372
7	799	14.2	915	21	Y78514
8	790	14.0	957	15	R56990
9	788.5	14.0	928	15	R56989
10	787.5	14.0	928	15	Y31746
11	777	13.8	826	20	Y31748

12	777	13.8	830	20	Y31747
13	716	12.7	849	20	Y24380
14	695.5	12.3	1938	17	W09255
15	692.5	12.3	1938	20	W33553
16	677	12.0	658	13	R23787
17	672.5	11.9	772	19	W49871
18	652.5	11.6	772	18	W34567
19	650	11.5	798	20	W35052
20	647.5	11.5	893	17	W09257
21	633	11.2	1250	21	Y1279
22	564	10.0	243	20	Y04387
23	532	9.4	931	20	Y27357
24	518	9.2	716	16	R71616
25	449.5	8.0	618	19	W5098
26	358	6.4	620	21	B05950
27	353	6.3	774	19	W35390
28	352.5	6.3	777	17	R92682
29	352.5	6.3	777	19	W3885
30	349	6.2	746	17	R92693
31	349	6.2	746	19	W3886
32	336.5	6.0	208	20	W81971
33	335	5.9	776	21	Y78516
34	334	5.9	750	10	P84419
35	320	5.7	745	10	P80615
36	319.5	5.7	726	20	W95636
37	319.5	5.7	726	21	Y78515
38	314.5	5.6	666	20	Y34991
39	307.5	5.5	766	20	Y00870
40	274.5	4.9	164	17	R88372
41	274.5	4.9	793	20	Y00869
42	267	4.7	818	20	Y17523
43	265.5	4.7	164	18	W34879
44	265	4.7	764	21	Y50819
45	264	4.7	618	20	Y37184

#### ALIGNMENTS

RESULT 1  
ID P82507 standard; protein: 1096 AA.  
XX  
AC P82507;  
XX  
DT 01-NOV-1990 (first entry)  
XX  
DE Pullulanase protein.  
XX  
KW Pullulanase; starch; alcohol prodn.  
XX  
OS Klebsiella aerogenes.  
XX  
PN JP63245676-A.  
XX  
PD 12-OCT-1988.  
XX  
PE 31-MAR-1987; 87JP-0078355.  
XX  
PR 31-MAR-1987; 87JP-0078355.  
XX  
PA (ELSD ) DENKI KAKAKU KOGYO KK; (SUNR) SUNTORY LTD.  
XX  
DR WPI: 1988-333488/47.  
XX  
N-PSDB: N81341.  
XX  
PT Gene encoding pullulanase - derived from recombinant plasmid pmp1 contg.  
XX  
PS Gene from Klebsiella genus.  
XX  
CC Disclosure; ; p; Japanese.  
CC The pullulanase protein cleaves alpha-1,6-glucoside bonds of starch and  
is effective in decomposition of branched starch. It is used in the



```

OY 232 -----TFTVTSQOVSMEFPHLSTIAAFKLPDANAVDELLOGETVAIAAEDGILISATQ 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 dkdggglpanvte---kfhrlrfsatfkapacltdvdsllkqclavaafsadqacrnalg 190
OY 286 VOTAGVLDADAEEAAEALSY-----GAQLADGVTFRVWAPTAQOVDVVVYVYADKRVIGSH 341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 lqlpgvldely-----sydqlgavfsehtsilylwaptaqavsaftf---kdpqgqe 240
OY 342 PMTR---DSASGAMSGMOGSDLKGAFAFRYAMTVYHPOSRKVEQYEVTPDYAHSLSSTNSEY 398
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 plqlvqllesngvswavayprlwegecyvyyeltvynhsrlrlekstaidpyaaglsadvtr 300
OY 399 SOVVDLNDLSALKPDGMDNL/T---MPHAQTKRADLAKMTIHESHIRDL/SAMDQTVPAELRCK 456
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 tlladlsetlkepgwenladekphllsp---sdlsilyelhlrdfsaydlvhpdlcrg 356
OY 457 YLALTAGSNMVOHLKTLTASAGTVHVELLPVFDLATVNEFSGKVADIQOPFRSLCEVNSA 516
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 ylatfsqdsagvnhleklsaaqltnvhlpsfqtaevdd--dk----- 397
OY 517 VKSSEFAGYCDSGSTVEEVLNQLKQSDQDNPOVALNTLVATQPSYMWGYPFHVTYPE 576
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 -kwwf-----vdekrfcltppdseeqgqitaitdeegynmgynpvlwgtprk 444
OY 577 GSYATDPGCTTRIKERFTWIAQIKODLGMNVIMDVYVYHFTNAAGPTDRTSLDKIVPWY 636
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 445 gsyadpmpgpcrllefkrmwqaln-r-iglrvaldvynhlnesspddsvldklvpyy 503
OY 637 QRLNLTSSVESATCCSAPREHMFARKLIADSLAVWTTDYKIDPFRRLMCTYHPRAOQL 696
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 504 lr-rindgalenstcvtndtaehfmerliddlkhwavnykvdgfrldlmhlmhltw 562
OY 697 SAMERIKALNPD-----IYFEGEGMDSNQ---SDRPEISQINLKGIGTIGTFSRRL 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 563 katmlglglsknlidvegssilylgegwdfgevanmargvasqnlsgtqsgfndir 622
OY 746 DSVRGCGPPDGDALRQNOGIGSGAGVLPNE---LASISDDOVRLHA-DLTRLGMAQLA 801
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 623 davlfgpfr--gpplq--qgyvtglslqpdhdhsgkanadtrmlavakdhlgvmaglnr 678
OY 802 DEVMIDKXGAANKKSEI-DYNGAPGAGYADPTEVVYVYSHKNOITMDISYKASQEAQL 860
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 679 dylfncdgaqkykgeevtylgytgvayqplctlnyvasahnetlfdlvsllkptyltv 738
OY 861 ATRVEMQAVSLTAVMAGOGIAFDQOGSELLRSKSFTRDSDGEMFERNRYSLODNNYV 920
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 739 decervnhatelalsqgipfifhagdelrlrkslrdsgynsgdwfnrlidfsynmngv 798
OY 921 GMPRISDDGSNTFVTRVAKENYATPG--EAELKQMTAFYQELT---ELRKSPLFTLGDS 976
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 799 glpdkhnesnwpil---krtlanpsykpkdknhllaavenfnllqlrfsyppflfrtsak 855
OY 977 AVMKRVDFENRTSSDQAGLLVMTVDGMAKGA-----SLDSRLDGLVAVTANAP-ESRTLN 1031
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 856 diedvrrlhmnyvswlpgllamsleog-hagapglsglqpkfyivllvngpcteklvn 914
OY 1032 -EFAGETQLASAIQOTAGENSLANGVOIADGTVTLPAWSVALELPO 1078
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 915 pllrkskslqhpvgatsgdvtvkesskypstgctflprkstavefr 962

```

```

RESULT 3
W81973
ID W81973 standard; Protein: 878 AA.
AC W81973;
XX 02-MAR-1999 (first entry)
XX Corn pullulanase protein fragment.
XX pullulanase; corn; transgenic plant; starch structure; isolation;
XX screening; developmental stage.
KW

```

```

XX OS Zea mays.
XX PN MO9850562-A1.
XX PD 12-NOV-1998.
XX PF 04-MAY-1998; 98MO-US09102.
XX PR 06-MAY-1997; 97US-0045723.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Broglie KE;
XX DR WPI: 1999-034728/03.
XX DR N-PSDB: V65118.
XX PT New nucleic fragments encoding corn pullulanase - useful to alter
XX pullulanase levels and thus starch structure in a host cell
XX Claim 1; Fig 1; 42pp: English.
XX PS This sequence represents a fragment of a novel Zea mays pullulanase
XX protein. This protein and the encoding nucleic acid fragments may be used
XX to create transgenic plants in which corn pullulanase is present in cells
XX at levels differing from the norm, or in cell types or developmental
XX stages where it is not normally found. In order to alter starch structure
XX in those cells. They may also be used to isolate cDNAs and genes encoding
XX homologous pullulanase from the same or other plant species, to amplify
XX those genes and to produce antibodies for screening of cDNA expression
XX libraries.
SQ Sequence 878 AA:
Query Match 25.2%; Score 1419.5; DB 20; Length 878;
Best Local Similarity 35.7%; Pred. No. 3,5e-99;
Matches 331; Conservative 172; Mismatches 340; Indels 83; Gaps 23;
OY 185 AEAHWDKNTLLMPGGQDKPIVLYLTYSHSSKVAADDEG-KFTDRYLAKLPTT--VSQOVS 241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 araywvtsklianlsdqktslfllyasrnaicmsqdkmkygskvelqpendgllpsvt 62
OY 242 MRPHLSYAAAFKLPDANAVDELLOGETVAIAAEDGILISATQVOTAGVLDADAEEAAE 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 qkprfissyrafripssvdatlvkqclavaafsdalgnrgdvglqlpyldmftfytg- 121
OY 302 ALSYGAQLADGVTFRVWAPTAQOVDVVVYVYADKRVIGSHPMTR---DSASGAMSGOG 357
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 -plglisseaavsmjlwaplaqdvsvsfydgp-----agplletvqlnelngvavtgp 174
OY 358 SDLKGAFAFRYAMTVYHPOSRKVEQYEVTPDYAHSLSSTNSEYSOVVDLNDLSALKPDGMDNL 417
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 rtwenryylyevlvynqctgnleklaadpyaaglsanstlrvlvdlnetlprlwdgl 234
OY 418 TWPHAOCTRAD-LAKMTIHESHIRDL/SAMDQTVPAELRKYTLAFLTAGSNMVOHLKTLISA 476
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 a---aekprldsfdsilsyelhirdseahdstvdcprfgfcftqdsvglshlkkled 291
OY 477 SGVTHVELLPVFDLATVNEFSDKVADIQOPFRSLCEVNSAVKSSERAGYCDSGSTVEEV- 535
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 agltvhlpsfqfgvdd-----lks-----nkwcvdele 322
OY 536 LNDLKO-SDSQDNPOVALNTLVAOTDSYMWGYPFHVYTPBEGSYATDPGCTTRIKERFT 594
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 lsklppgsdlq-----gaatvaideedpymwgnpwwgvkgsyasnpqgparlleyr1 377
OY 595 MIOAIRKODLGMNVIMDVYVYHFTNAAGPTDRTSVLDKIVPWYQRLNETTGSVATCCSD 654
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 378 mwqaln-r-iglrvaldvynhllyssgpfaltsvldklvpyylyr-rdsngqtensaavn 435
OY 655 SADEHMFARKLIADSLAVWTTDYKIDPFRRLMCTYHPRKAOILSAMERIKALNPD----- 708

```

```

Db      436 taeshfnvdlivddllnwvnykvdgfrldlmglnhkkcmiraksalqsltldehvgdg 495
OY      709 --IYFESEGMDSN--OSDRFEIASQINLKGTGTFSDRLRDSVRGGPDSGDALRON 763
Db      496 skilygyevnfgvaengrfgngsqikmsgtgysfndtrldangspfr--gnplq-- 551
OY      764 OGIGSGAGVLPNELASISDDOVR---HLADLTRLGAGNGLADPVMIDKDGAAKSGSEI- 818
Db      552 qglsfqlflepngfygnetelrtllaalyadhqlgldkdyvvlshgsearkgselr 611
OY      819 DYNGAPGAYADPTEVENVYVSKHNDOTLMDMISYKASQEADLARTVMQAVSLATVMLGQ 878
Db      612 tfdgsprvygsspletlnygsahndelctldlstkpmdisidercrinhstsmalsq 671
OY      879 GIAFDQGSSELRKSFTRDSYDSGDFNRYDYSLODNNYVGN-PRISDGSNYEYTR 937
Db      672 glpfthgdeltrksaidrdsydsdgdfnkldfycelnmgvgjpprekegswplmkpr 731
OY      938 VKEMVATPGCAELKQMTAFQOELTELKSSPLFTLGDSAVMKRVDRNTGSDQOAGLLV 997
Db      732 lempstfkpakhdliaaldkldlklrfsplflrltaadlvgrvfhncgpalvpgvly 791
OY      998 MFDVDDGKAC---ASLDSRLDGLVVALNAAP--ESRTLNEPACFTQLSALQOTAGENSU 1052
Db      792 msledarndrhmagidctscvvtvfnvcpvevsteipdlastrlqlhnpqvnsdala 851
OY      1053 ANGVQIADGTVTLPANSAVLELPQ 1078
Db      852 rgsaydlatgltfcvprkrtaaavfepr 877

RESULT 4
W34878 ID W34878 standard: Protein: 558 AA.
AC      W34878;
DT      25-MAR-1998 (first entry)
DE      2. mays debranching enzyme.
KW      debranching enzyme; maize; transgenic plant; starch; food industry;
        industrial product.
OS      Zea mays.
FH      Key
FT      Protein
XX      Location/Qualifiers
        note="partial sequence of debranching enzyme"
PN      DE19608918-A1.
XX      11-SEP-1997.
PD      07-MAR-1996; 96DE-1008918.
PF      07-MAR-1996; 96DE-1008918.
PR      07-MAR-1996; 96DE-1008918.
XX      (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
PI      Emmermann M, Kossmann J, Willmitzer L;
XX      WPI; 1997-458572/43.
DR      N-PSDB; T93875.
XX      DNA encoding maize de-branching enzyme - useful for production of
XX      starch in transgenic plants
XX      Claim 1, Page 19-21; 13pp; German.
XX      This partial sequence represents a novel debranching enzyme isolated from
XX      Zea mays (maize). This enzyme has been used in the production of

```

```

CC      transgenic plants in which the nucleic acid is stably integrated into the
CC      host cell's genome and expressed. Such an enzyme can be used in the
CC      production of starch which is useful in the food industry or in the
CC      production of industrial products e.g. paper, adhesives, textiles,
CC      textile care products, building materials, soil stabilisers,
CC      agrochemicals, pharmaceutical products, cosmetics, coal briquettes,
CC      flocculants for ore and coal slurries, foundry sand binders, rubber
CC      goods, leather substitutes and polymers.
XX      Sequence 558 AA:
SQ
Query Match 17.7%; Score 997.5; DB 18; Length 558;
Best Local Similarity 39.0%; Pred. No. 2, 2e-67;
Matches 218; Conservative 110; Mismatches 198; Indels 33; Gaps 12;
OY      542 SDSQDNPOVALNLTLVQTSYVNGVDPFHYTVEGSAVDDETTIKERFMIAQIKQ 601
Db      10 sdly-----qaalvaldeedpymgynpvygypkgsyasnbdpsrlleyrlmwgalnr 64
OY      602 DLGNVIMDVVYVNTNNAAGPTDRTSVLDKIVPMYYQRLNETTGSVESATCCSDAPHRM 661
Db      65 -lglrvmdvnyhnllysgpfaltsvldkivpgyll-rdsngtensaavmntasehfm 122
OY      662 FAKLIADSLAVWTTDKRIDGFRFDLMGYHPRKAQILSAMERIKALNP-----IYFG 713
Db      123 vdrllvddllnwvnykvdgfrldlmglnhkkcmiraksalqsltldehvgdgakilyg 182
OY      714 EGMDSN---OSDRFEIASQINLKGTGTFSDRLRDSVRGGPDSGDALRONIGSGA 770
Db      183 egvnyfgevaengrfgngsqikmsgtgysfndtrldangspfr--gnplq--qglsfql 238
OY      771 GVLNPNELASISDDOVR---HLADLTRLGAGNGLADPVMIDKDGAAKSGSEI-DYNGAP 825
Db      239 flpangfygnetelrtllaalyadhqlgldkdyvvlshgsearkgselrtfdgspr 298
OY      826 GYAADPTEVENVYVSKHNDOTLMDMISYKASQEADLARTVMQAVSLATVMLGQIAFDQO 885
Db      299 gyaaspictlnygsahndelctldlstkpmdisidercrinhstsmalsqldpfha 358
OY      886 GSELRKSFTRDSYDSGDFNRYDYSLODNNYVGN-PRISDGSNYEYTRKEVAT 944
Db      359 gdeltrksaidrdsydsdgdfnkldfycelnmgvgjpprekegswplmkprlmpsfk 418
OY      945 PGEAEELKQMTAFYQOELTELKSSPLFTLGDSAVMKRVDRNTGSDQOAGLLVTVDDGM 1004
Db      419 pakhdllaaaldkldlklrfsplflrltaadlvgrvfhncgpalvpgvlymsledar 478
OY      1005 KAG---ASLDSRLDGLVVALNAAP--ESRTLNEPACFTQLSALQOTAGENSILANGVIA 1059
Db      479 ndrhmagidctscvvtvfnvcpvevsteipdlastrlqlhnpqvnsdalarqsydt 538
OY      1060 ADGTVTLPANSAVLELPQ 1078
Db      539 atgrfvcprkrtaaavfepr 557

RESULT 5
ID      Y78513 standard: protein: 862 AA.
AC      Y78513;
DT      05-MAY-2000 (first entry)
DE      Pullulanase (pulB) amino acid sequence.
XX      Pullulanase; pulB; debranching enzyme; amylopectin; thermostable; corn;
XX      genetically engineered variant; wheat; potato; wheat; manioc; rice;
XX      starch; sugar production; high fructose syrup; high maltose syrup;
XX      maltodextrin.
OS      Bacillus acidopullulyticus.

```



XX WPI, 1998-003022/01.  
 DR N-PSDB: T97233.  
 XX  
 PT DNA encoding a trimmed enzyme - used in the production of a trimmed  
 XX enzyme with wide substrate specificity  
 PS  
 PS Claim 2; Page 9-10; 15pp: Japanese.  
 XX  
 CC The present sequence represents a trimmed enzyme isolated from  
 CC *Bacillus* sp. APC-9603. The present specification also describes:  
 CC (1) a vector containing the DNA encoding a trimmed enzyme; (2) a  
 CC recombinant microbe transformed by the DNA; and (3) preparation of  
 CC a trimmed enzyme in which the microbe of (2) is cultured and the  
 CC trimmed enzyme is collected from the culture. The method can prepare  
 CC a trimmed enzyme of wide substrate specificity having heat resistance  
 CC and being active in acid region efficiently.  
 XX  
 SQ Sequence 922 AA:

Query Match 14.6%; Score 820; DB 19; Length 922;  
 Best Local Similarity 24.9%; Pred. No. 1.6e-53;

Matches 277; Conservative 155; Mismatches 353; Indels 326; Gaps 36;

QY 79 DYSSKNLYLNNETCDLALSPVADMDVSTTPSGSKYIGPYWVPLKESGCIWVYRD-137  
 DB 17 dygswnslwmwpeg-----gdgnynf--ngtdsgyelanvslpyspakvylvtq 65  
 QY 138 -----GTDKLIDSDLRVAFGDFTRTVSVYAGNSAVYDSRDAFRALFVLAELAHMVD-191  
 DB 66 dwakdvsgdrlvldls-----khevwlyvqnsqilfynekdaedaa-epavsnaylda 116  
 QY 192 KNTLL-----  
 DB 117 pnkvlvksqpfcltlegasgftvhdclangdlpvtkvnankevdeallagltfqlhleg 176  
 QY 197 --W-----PGGDKPIVRLYSHSSKVAADGEGKFTDRILKL-231  
 DB 177 sdwaphshsqllkvndhlygfsgelipgsyqykvvalindwssyspdcn-----lnl 228  
 QY 232 TPTTVSQVSMRPHLSYAFKLPDNNANVDELIG-----ETVAIAAEEDGILISATOV-286  
 DB 229 tvpdsghavhtsyvp-athavvdtlmpnganlprldsgskldvltvclgenpdshtlsl 287  
 QY 287 QTAC-----VLD-DAYAEALSYCAQLADCGVTFRRVAPTAQOVVVVSADK-335  
 DB 288 qldgfkqgrvlpnvlfdsgyyysgedl--gnytkkattfkwapstckvnllynn-342  
 QY 336 KVISIHMT---RDSASGAWMGGSDLKGAFTYAMTVIHPSRKVEQVETVPIAHSL-392  
 DB 343 kbaagalkctvpmkasghvswvtpvgnlgnwylyevltvgqstrlav-----dpyatai 396  
 QY 393 STNSEYSQVVDLINDSALKPDGMDNLTMPHIAQTKADLAKMTIHESIIRDLSAMQTPPAE-452  
 DB 397 apngltgmwvl--akcnprgwk-----dkhmlpknldevelymhrvrdfei-dsngmt 449  
 QY 453 LRGKYLAL---TAGDSNMVQHLKTLASGVTHVELLPVFDLATVNEPSDKVADIQPFS-508  
 DB 450 nkgylyaltelkqgpenvklygvdalkqlglthvqlpvnafnsvde-----496  
 QY 509 RLCEVNSAVKSEFEAGYCDGSGTVEEVLNQLKQSDSDNPQVQALNTLVNQT--SYNWG-566  
 DB 497 -----tdptqymw 505  
 QY 567 YDPFHYVPEGSYATDEGTRRIKEFRTMIOAI-KODLGMNVINDVYVNHFNAGPFDRT-625  
 DB 506 ydprnyvnpqgyatdangtrcltkefkemvlslnrhlygn--mdvynhbfafqisd--561  
 QY 626 SVLDKIVPWYIQRLETTGVSYSATWCSDSAPBHRMFAKLADSLAVWTTDYKIDGFRFD-685  
 DB 562 --tdklypvyvctddagvnylunsgltgnevnaerpmvqkflldalkywnveyhldgrfdd 619

QY 686 LMGYNPKAQLISAMERIKALNPDIYFEGGWDNSQS--DREELASQINLKGITGTFSDR-743  
 DB 620 lmalllgkdtmakaeqelhaldrpialygerwtggtcalprcdqlltkgvqkmyvavfndn 679  
 QY 744 LRDSVRGCGPDPDSCDALRQNOGICSGAGVLPNELASLSDDQVRLADLTRUGMAGNLADF-803  
 DB 680 lrrgl-dgnvfda-----ssqgfata-----tgltavlkkygeylndf 718  
 QY 804 VMIDKGAARKKSEIDYNGAPGCGYAADPTEVNVNSKHDQTLMDMISYKASQEADLATR-863  
 DB 719 -----tsapgetlioytshdyltwkklagpnndse-adr 753  
 QY 864 VRMOAVLATVMLGQIAFDQGSSELLRSKSFTRDSGDMENRVDYSQDNNVNGMP-923  
 DB 754 lkmdeiaqavvvtsgyprfmgqgemlrfkqgnsynagdavvefdvs-----802  
 QY 924 RISDDGSNVEITFRVKEMVATPGAEELKQMTAFQELTELKSSPLFTLDDGSVMKRYD-983  
 DB 803 -----rkaqsdvlnyysglhlhrlahpalmrltanqlkhehq 840  
 QY 984 FRNTGSDQAGLLVMFVDDGMKAGASLDSRLDGLVVAINAPESRTLNEFAGETLQLSAI-1043  
 DB 841 fidespdr-----tvayeltahand-kwnlvtvlynpuktaetvnlpsqk-walnat 890  
 QY 1044 QQTAGENSLANGVOIADGVTLTPAMSVAVL-1074  
 DB 891 ngkigeslsh-----aeghvqyvgismml 916  
 RESULT 7  
 ID Y78514 standard; protein; 915 AA.  
 AC Y78514;  
 XX  
 DT 05-MAY-2000 (first entry)  
 XX  
 DE Pullulanase amino acid sequence.  
 XX  
 KW Pullulanase; debranching enzyme; amylopectin; thermostable; corn;  
 KW genetically engineered variant; wheat; potato; wheat; manioc; rice;  
 KW starch; sugar production; high fructose syrup; high maltose syrup;  
 XX maltodextrin.  
 OS  
 OS *Bacillus* *deramificans*.  
 PN MO200001796-A2.  
 PD 13-JAN-2000.  
 XX  
 PE 02-JUL-1999; 99WO-DK00381.  
 XX  
 PR 02-JUL-1998; 98DK-0000868.  
 XX  
 PA (NOVO ) NOVO-NORDISK AS.  
 XX  
 PI Bisgard-Frantzen H, Svendsen A;  
 DR WPI: 2000-160767/14.  
 XX  
 PT Variant bacterial pullulanases and isomylases having, e.g. increased  
 PT thermostability, used for converting starch from potatoes into high  
 PT fructose syrup  
 PS  
 PS Claim 29; Fig 3; 116pp: English.  
 XX  
 CC This sequence represents a pullulanase amino acid sequence. Pullulanases  
 CC are debranching enzymes which can attack amylopectin. This sequence  
 CC represents a parent enzyme from which genetically engineered variant  
 CC debranching enzymes can be created. The genetically engineered enzyme has  
 CC an improved thermostability at a pH in the range of 4-6 compared to the  
 CC parent enzyme. The modified enzyme may alternatively or also have an  
 CC increased activity towards amylopectin and or glycogen compared to the



CC parent enzyme. The enzyme variants are useful for converting starch to one or more sugars. This is useful when starches such as corn, potato, wheat, CC manioc and rice starch are used as starting materials in commercial CC large scale production of sugars, such as high fructose syrup, high CC maltose syrup, maltodextrins, amylose, G4-G6 oligosaccharides and other CC carbohydrates such as fat replacers. The thermostable debranching enzymes CC of the invention make it possible to perform liquefaction and debranching CC simultaneously before the saccharification step. By debranching using the CC thermostable debranching enzyme variants during liquefaction together CC with the action of an alpha-amylase, the formation of panose precursors CC is reduced, thereby reducing the panose content in the final product and CC increasing the overall saccharification yield.

XX Sequence 915 AA:

Query Match 14.2%: Score 799; DB 21; Length 915;  
Best local similarity 25.0%: Pred. No. 6.1e-52;  
Matches 278; Conservative 144; Mismatches 353; Indels 338; Gaps 38;

QY 77 AADYSSKMLYLMN-----NETCDALAPVADWMDVSTTPSGSDKYGPYWIPLNK 126  
D 15 agdygpwalmwmpkdggaeydfnqpadsgfivas--adipgnpq----- 58  
QY 127 ESGCLNIVRD-----GTFKLIDSLRVAFGFTDRIVSYIAGNSAVYDSRADAFRAA 179  
D 59 -----vgllivrtqdwtkdvsadylids-----kgnevvlveqnsqiflynekaadea 106  
QY 180 FGVALAEAHMWVDKNTLMPGQDKRIVRLYYSHSSKVAADGSK--FT--DRYIKLTPT 234  
D 107 -kpravnsmlasngvl-----vkl-----sqpllgeasgfivhdtdankdipv 151  
QY 235 TVSQGVSM-----RFPHL--SSYAFKLPDN----- 258  
D 152 tsykasaslgdvtavlaqtfqhlfggsdwa----pdnhsllkvtmnllycfsqdlpegn 207  
QY 259 -----ANVDLLQGETVAIAAEDGILISATVOYQVAGLDDAVA----- 297  
D 208 yqykvamshsaqrvtsqavsvwpgrpylgaetwdqlygnvalysqsvkldlvtlvegedp 267  
QY 298 EAAEALS-----YGAQLADGQVTFYRWAPTAQOVDVY 329  
D 268 dshslslsgtdgqyakqylyprnvlnssqyysgddlgnlytcattkfwapstcqvnyl 327  
QY 330 VY-SADKRVYIGSHPTRDASAGAMSGWOGSDLKGAFFRYRATVYHPQSRKVEQYEYTDTPY 388  
D 328 lydsatgsvtklvpmta--sgbgvweatvqnleawymyvevcgsstrlav-----dpy 380  
QY 389 AHSLSSTSEFSYVVDLNSALKRPDGMNLTMHAQKTKADLAKMTIHESHKIDLSAMDOT 448  
D 381 atalpngrlgrmlyvd1--aktdragwns---dkhltpknledevlyemdvdfsl-dpn 433  
QY 449 VPAELRGKYLAL-----TAGDSNMVOHLKTLASGVTHVELLPVFLATVNEFSDDKVAIDQ 504  
D 434 sgmkhkygylattekgtvgprdhvktgldsklgilthvqimpyfasnsyde--tdprtd-- 490  
QY 505 QPFSRLCEVNSAVKSSSEFAGYCDSSGTYEEVLNOLKQSDSDQNPQVALNTLVAQTSYN 564  
D 491 -----n 491  
QY 565 WGYDFPHYTVPEGSYATDEGTTTRIKERFTMTQAT-KODLGMNVMIDVYVNHNTMAAGPRD 623  
D 492 wgydfprnydvpegyatnangnariketkewlslshrehlygn--mdvynhlftafqjald 549  
QY 624 RTSVADKIVPWYQRLNETTGSVESATCCSDSAPHEHMAKLIADSLAWTMDYKIDGR 683  
D 550 -----fdklvpeyyrttdagnytnsgsgneiaaerpmwqkfilidslkywvneynhidgr 605  
QY 664 FDLMGYHRAQILSAMERIKALNPDIYFFEGGMDSNOS--DRFELASQJNLKGTGIGTS 741  
D 606 fdlmallgkdtmskaaselhainpjalypgpwtlgtalppddqlltkgaqgmgyavah 665  
QY 742 DRLRDSVRCGGCFPDSGDALRQNOIGSGAGVLPNELASLSDQVRLADLTRLGMAGNLA 801

DB 666 dhlrnl-dgnvfds-----saqfataga-----tgldaknryegsnh 704  
QY 802 DEVMIDKDAAKKSGSEIDYNGAPGYADPTEVYVNSKHDNQTLDMDSYKASQADIA 861  
D 705 df-----tspgetlnvthdhvtylvdcklalsnpsnse-a 739  
QY 862 TVWRQAVSLATVMLGQGIAPDQGSSELLRSKSPTRDSYDSGDWFMNRVDYSLQDNNYNG 921  
D 740 dtkmdelagavmtesgyprfmggsemlrltkgndnsynagdavnefdws----- 790  
QY 922 MRISDGSNTEVITVKKEMVATPGEAELEKQMTAFYQELTELKSSPLTFLDGSAYMKR 981  
D 791 -----rkaypdyvinyqgllhrlldhpatrmtaneish 826  
QY 982 VDFRNTGSDQAGLVMYTVDDGKAKGASDSRLDGLVVAIINAPESRTINEFAGETLRS 1041  
D 827 lqflnspevtvayeltdhvn-----kdkwgnllyvynpnkvtvalinlpsqk-wain 876  
QY 1042 AIQQTAGENSLANGVOIADGTVTLPAWSVAVL 1074  
D 877 atsgkvgestlqg-----aegsvqvpjismml 904

RESULT 8  
R56990  
ID R56990 standard; Protein; 957 AA.  
XX  
AC R56990:  
XX  
DT 21-FEB-1995 (first entry)  
XX  
DE Bacillus deramificans pullulanase precursor.  
XX  
KW Pullulanase; recombinant production; pullulan; amylopectin;  
KW hydrolysis; alpha-1,6-glycosidic bond; starch saccharification.  
XX  
OS Bacillus deramificans T 89.117D (LMG P-13056).  
XX  
FH Key Location/Qualifiers  
FT MISC-difference 163  
FT MISC-difference 164  
FT MISC-difference 164 /note= "corresponds to GCN codon"  
FT MISC-difference 620 /note= "corresponds to NNA codon"  
FT MISC-difference 621 /note= "corresponds to CAN codon"  
FT MISC-difference 621 /note= "corresponds to GCN codon"  
FT MISC-difference 812 /note= "corresponds to ANA codon"  
XX  
PN EP605040-A.  
XX  
PD 06-JUL-1994.  
XX  
XX 20-DEC-1993; 93EP-0203593.  
XX  
XX 28-DEC-1992; 92BE-0001156.  
XX 15-JUL-1993; 93BE-0000744.  
XX 19-NOV-1993; 93BE-0001278.  
XX  
XX (SOLV ) SOLVAY SA.  
XX  
XX Amory A, Deweer P;  
XX  
XX WPI: 1994-210291/26.  
XX N-PSDB: Q68699.  
XX  
XX New pullulanase from Bacillus deramificans - for starch  
XX saccharification etc., with good stability over wide temp. and pH  
XX ranges, also related DNA vectors, transformed cells etc.  
XX  
XX Claim 4; Fig 5; 61pp; French.



CC enzyme can be used (with glucoamylase) for saccharification of  
 CC starch. Typical applications include use as anti-staling additives  
 CC in bread making and brewing, in preparation of low calorie foods  
 CC and to clarify fruit juices.

SO Sequence 928 AA;

Query Match 14.0%; Score 788.5; DB 15; Length 928;  
 Best Local Similarity 25.4%; Pred. No. 3.9e-51;  
 Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;

QY 77 AADYSSKULYLN-----NETCALSAAPA-----DW-NDVST 108  
 Db 15 agdyqpsalmmprkdggaeydfnqpadsgvasadilpnpbgylivrtgdkvda 73  
 QY 109 TPSSGDXY-----GPVWVPLNKE-----SGCINIVRGTDLKIDSLRV 149  
 Db 74 ----adryldlstkneevlvegnqglfynekdaedaakpavsnayldsnqylvklspql 129  
 QY 150 AFGD-----FT-DRTVSYIAGNSAVYDSR-----ADAFRAFGVALAEHWDKN- 193  
 Db 130 tlgeqxsqftvhdtdanldipvtsvkdaslgdvtavlagtqlhlfq-----gsdwapdh 185  
 QY 194 -TLMPGQDKRYRLVYSHSKVADSGKF-----TDRLKLPPTVSQOVSNRP- 245  
 Db 186 stll-----Kkvtmlyfsgdlp-----egnyqvvalndsw--nmpsyndhiltvpa 234  
 QY 246 ---HL-----SSYAFKLPDMANDELLOCT-----VALAAEDGLILATQVOTNG 290  
 Db 235 gqahvtfyslsthavtltmnpnad--lqvsgvktldvltvlgdpshtlsiqdg 292  
 QY 291 -----VDDA-VAEAEALSYGAOLADGVTFRWAPPAQOVVVV--SADKRYI 338  
 Db 293 yqakvylprnvlnsqyyysgddl--gntylqkaltfkfwapsqgnvllidsatgvt 350  
 QY 339 GSHPTKDSASGMSWGSDLKGAFTRYAMTVHPQSKVQEVETDPAHSLSTNSEY 398  
 Db 351 klvmta--sgbhvveatvqnlennvymyevtgsstrlav-----dpyataiapngtr 403  
 QY 399 SOVVDLNSALKPRDGMNLTMPHAKTKADLAKMTIHESHINDLSAMOTVFAELRGKL 458  
 Db 404 gmlvdl--aktidragwns---dkhlprknldevilyendvdfsl--dpsngmkngkyl 456  
 QY 459 AL-----TAGDSNMVQHLKTLASAGTHVELLPFDLATVNEFSRVDADIQPFSLCEVN 514  
 Db 457 altegtgprdnvktgtslkgjltlvqimpvfasnsvde--tdrptqd----- 503  
 QY 515 SAVKSSERAGYCDSGSYVEEVLNQLKSDSDPNPOVALNTLVAOTDSYNMGYDPEPHYV 574  
 Db 504 -----nwgydprnydv 514  
 QY 575 PEGSATDPEGTRIKFERTMIQAI--KODLGMNVIMDVVYNNFTNAGAPTRDTSVLDKYV 633  
 Db 515 pegqatnangnarikelkewmvlshrenllyvn--mdvvynltlftqlsd---tdklyp 568  
 QY 634 WYQRLNETTGSVESATCCSDSAPHRMFA-----KLADLAWTDTYKIDGFEFDLM 687  
 Db 569 eyyyt-----ummgyliprtqvltemkixaeerpmwqkfidslkyvnehyhldgffidm 622  
 QY 668 GYHPRAQLISAMERIKALNPDIYFFEGGWDNSQS--DREELASQILNKGITGTFESDLR 745  
 Db 623 allgtdmkaeselhahpajalygepwtgtsalpdqlltkgqkmyavavfndlr 682  
 QY 746 DSVRGCGPDSGDALRONCIGSAGVLPNELASLSDQVRLADLTRIGMGNLADFWM 805  
 Db 683 nal--dgnvfd-----saqgfatg-----tgltdaikngvsgslndf-- 719  
 QY 806 IDKDAARKSEIDYNGAARGAADPEVYVYVSKHNDOTLMDMLSYKASQADLATRYR 865  
 Db 720 -----tspsgetinytshndytlndkklaslmpnase--adrlk 756  
 QY 866 MOAVSLATVMLGQIAFDQGSSELRKSKFTRDSDSGDMENKVDYSLDDNNYVNGMPRI 925

Db 757 mdelqavvmtsqyplfmqgsemlrxsgndnsynagdavnefdws----- 803  
 QY 926 SDDGSNVEYITRVKEMVAPGEALKQMTAFYQELTELKSSPLFDGSAVMKRVDR 985  
 Db 804 -----rxkqypdvfnysgllhldhpafrmtlaneishqfl 843  
 QY 986 NTGSDQAGLVMYVDGKAGASLDSRLDGLVWINAPESRTLNEFAGETLOLSAIOO 1045  
 Db 844 nspentvayelcdhn-----kdkwgnllivynpukvtatlnlpsgk-walnatg 893  
 QY 1046 TAGENSLANGVOIADGVTYTLPAWSYAVL 1074  
 Db 894 kvgestlqg-----aegsvqvpjismml 917  
 RESULT 10  
 Y31746  
 ID Y31746 standard; Protein: 928 AA.  
 AC Y31746;  
 XX  
 DT 22-NOV-1999 (first entry)  
 DE Bacillus deramificans mature pullulanase.  
 KW Pullulanase; enzyme engineering; starch; saccharification;  
 KW straight chain amylase; alpha-dextrin 6-glucano-hydrolase;  
 KW alpha-1,6-glucosidase.  
 OS Bacillus deramificans.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 134 /note= "encoded by GCN"  
 FT Misc-difference 135 /note= "encoded by NNA"  
 FT Misc-difference 591 /note= "encoded by CAN"  
 FT Misc-difference 592 /note= "encoded by GCN"  
 FT Misc-difference 783 /note= "encoded by ANA"  
 FT Protein /note= "99.928  
 FT Protein /note= "N-truncated pullulanase used in claimed  
 FT Protein 103.928 /note= "composition of Claim 31 (see Y31747)"  
 FT Protein /note= "N-truncated pullulanase used in claimed  
 FT Protein composition of Claim 32 (see Y31748)"  
 PN WO9945124-A2.  
 XX  
 PD 10-SEP-1999.  
 XX  
 PF 03-MAR-1999; 99WO-0504627.  
 XX  
 PR 04-MAR-1998; 98US-0034630.  
 XX  
 PA (GENV ) GENENCOR INT INC.  
 XX  
 PI Miller BS, Shetty JK;  
 XX  
 DR WPI: 1999-540851/45.  
 DR N-PSDB: X87941.  
 XX  
 PT New modified pullulanase for saccharification of aqueous liquefied  
 PT starch  
 XX  
 PS Claim 31; Fig 1A-E; 49pp; English.  
 CC The present sequence represents Bacillus deramificans strain  
 CC 189.117D mature pullulanase (EC-3.2.1.41). This enzyme has a pI of  
 CC 4.1-4.5, has a half-life of about 55 hr at 60 deg C and is capable









FT	Active-site	/note= "C-terminal alkaline pullulanase"
FT		1396..1481
FT		/note= "pullulanase active centre"
FT	Region	1396..1401
FT		/label= Region-I
FT		/note= "pullulanase active centre region I"
FT	Region	1460..1468
FT		/label= Region-II
FT		/note= "pullulanase active centre region II"
FT	Region	1493..1496
FT		/label= Region-III
FT		/note= "pullulanase active centre region III"
FT	Region	1576..1581
FT		/label= Region-IV
FT		/note= "pullulanase active centre region IV"
XX		
PN	W09635794-A1.	
PD	14-NOV-1996.	
PP	10-MAY-1996;	96WO-JP01243.
PR	10-MAY-1995;	95JP-0111547.
PA	(KAOS ) KAO CORP.	
PI	Ara K, Hatada Y, Igarashi K, Ito S, Kawai S, Ozaki K;	
DR	WPI: 1996-518682/51.	
DR	N-PSDB; T47878.	
PT	DNA encoding alkaline pullulanase having alkaline alpha-amylase	
PT	activity - useful for improving the efficacy of dish washing	
PT	detergents and detergents for clothes, esp. on starch soils	
PS	Disclosure: Page 42-54; 75pp; English.	
XX		
CC	Amylopullulanase (W09255) of Bacillus sp. isolate KSM-Ap1378 has an	
CC	N-terminal alkaline alpha-amylase moiety (see also W09256) and a	
CC	C-terminal alkaline pullulanase moiety (see also W09257). It is	
CC	the product of an open reading frame in a DNA fragment (T47878)	
CC	isolated using shotgun cloning and PCR (see also T47879-84). The	
CC	Amylopullulanase, alpha-amylase and pullulanase can each be	
CC	expressed in transformed host cells using vectors contg. the	
CC	appropriate gene fragments. The enzymes are useful as additives	
CC	in dish-washing and clothes detergent compns.	
XX		
Sequence	1938 AA;	

QY	292	-LDDVAEEAALATSYGQOL--ADGCVFERVWAPPAQAOVDVVVSA--DKVTSIHMTFRS	347
Db	1134	lIdemaydydk--lsgelInedgtatlckwspkacnsvslydkvdqaevalIdemaykd	1190
QY	348	ASGAMSWO-----GGSDLKGAFAFYRVAMTVYVHNPSKRVBOEYV-----DPVAHSLSY-NS	396
Db	1191	-rywvsvyLtkdntglIdslkgyyhyheIt-----hgdtvnlaldrpykmaawm	1239
QY	397	E-----YSQVVDNDLSALKPDGHDNLITMHAQTKADLAKKTIHESHIRDSLSANDQYVP	450
Db	1240	eagdkvvgkaalvdl--gsIgre-ldpadiIpgfekreclt----IyevhvrdfTs-dpnlIg	1291
QY	451	AEIKGKYLALTAGGSNNVQHLKTLTASGAVYHNVELLPVFLDALVYNEPSDKVADIQOPFSRL	510
Db	1292	edlkaqfgltf--astvkeIdyIqelgyvchIdllpmyyIefeneI-----esagem	1339
QY	511	CEVNSAVKSEFACVCDGSGSTVEEVLNQLKQSDQNPQVALNLTVALQOTDSYNNQGDPE	570
Db	1340	Ie-----yaatIg-----Iynnyghp	1356
QY	571	HYTPEGSATDPREG--TTRIKERFTMQL--KODLGNNYIMDYVYVHNTAAAPTDRISYL	628
Db	1357	nyfslsgmynscpdcpeIrltkefnlInelhkrdmg--vvladvInh-----taqvhlf	1408
QY	629	DKIYPMYQRLNETFGSVESATCCSDSAPRHNFAKLIDLSLAWMTTDYKIDGFRFDLGC	688
Db	1409	edlvpnyhfm-daddgprtsfgggrrlgltchemsrIyldslkhwydekvgdfIdmg	1467
QY	669	YHRAQILSAWERIKALNPDIYFECEGW--DSNOSDREIASQILNKGCT-GIGTFSRL	744
Db	1468	dhdeesIqladeakInpIvlmIgwefIagdegepyaqadqwmgyIteavgfsdeI	1527
QY	745	RDSVRGGGPRFSSGALNQNQIGSGAGVLPNELASLSDDQVRLIADLTILGAGNLADFY	804
Db	1528	rnelk-----sgsg-----segqrflIt-----	1545
QY	805	MIDDKGAARKSEI--DYNGAPCGYAD--PTEVYVYVSKHNDQILMDIMSYKASQBADLA	861
Db	1546	-----ggaevnvqglfndlkqphImedqgdgvdyvylaeahndltIydlavagIskkdpela	1600
QY	862	-----TRVRQAVSLATVYMLGGIATFDQOGSELLRKS-----	894
Db	1601	endleIhrlfI--vgnamvIteqgtaflIhaqgefgtrtkqvrapateapaykayltcdag	1657
QY	895	-----FTRSYSGCMFNRVQVSLSDNNYVNGMRISDDGSNVEVIRRVKEMVATPE	947
Db	1658	npfyvpyrlIhnsydsdIlhfrfdek-----atdaeKpyvnnvtIdya-----	1701
QY	948	AELKOMATFYQELTELKSSPLFTLGDGSAVMKRVDFRNTGSDQAGLILMTVDDGMKAG	1007
Db	1702	-----gllelrssdaflIgrrelvdanvcmvapeIkegdIyv-----ay	1742
QY	1008	ASLDSRLDLGVAINAAPESRFTLNERFGEFLQLSAIQIQTAGNSLAN-----GVQIADG	1062
Db	1743	rsvstagaeyyctfnadtsrcltI--Iggdltegvvvvdaeeanvagvapeagfelteeg	1800
QY	1063	TVTLPAWSVAVLELPOGEAOG	1083
Db	1801	-tlleplttvvvI--gegeg	1818
RESULT	15		
ID	W73553		
XX	W73553	standard; protein; 1938 AA.	
XX	W73553;		
XX	08-MAR-1999	(first entry)	
XX	Full length Pullulanase protein.		
DE			
KW	Pullulanase; variant; bleach-containing detergent; oxidation resistant		



```
XX  Bacillus sp.
OS
XX  Key Location/Qualifiers
FH  Misc-difference 121..137
FT  /note= "nucleotides encoding these residues are not
FT  present in the coding sequence given in the
FT  specification"
FT  Misc-difference 1050..1065
FT  /note= "nucleotides encoding these residues are not
FT  present in the coding sequence given in the
FT  specification"
XX  JP10327868-A.
XX  15-DEC-1998.
XX  30-MAY-1997; 97JP-0141596.
XX  30-MAY-1997; 97JP-0141596.
XX  (KAOS ) KAO CORP.
XX  WPI: 1999-099031/09.
XX  N-PSDB: V62881.
XX  New mutant pullulanase - useful in bleach-containing detergents
XX  Claim 13; Page 10-17; 19pp; Japanese.
XX  This sequence represents the Bacillus pullulanase protein. The invention
XX  relates to a mutant pullulanase in which at least one methionine residue
XX  in the pullulanase is deleted or replaced by an optional amino acid
XX  residue other than Cys and Met and/or at least one alanine residue in the
XX  amino acid residues constituting pullulanase is deleted or replaced by an
XX  optional amino acid residue other than Ala. The mutant pullulanase is
XX  useful in bleach-containing detergents. The pullulanase is highly
XX  resistant to oxidation.
XX  Sequence 1938 AA;
SO
Query Match 12.3%; Score 692.5; DB 20; Length 1938;
Best Local Similarity 24.2%; Pred. No. 2.6e-43;
Matches 281; Conservative 168; Mismatches 387; Indels 325; Gaps 51;
OY 11 LGSVL--LSCDNGSSSSSGNDPTPDNDVVRLLPDVAAPGEAVTVAVENQAVHLWD 68
DB 895 lglfvlntngkdgdgkavelfspdl--nelwlkgsdevfliepvdlpantvrlhy-- 950
OY 69 IAGITSSAADYSSKNLYLNNETCDALSAFVADWMDVSTPSCGSDKYGYPVWIPLNKES 128
DB 951 -----ernadvegglmwed-----vespsdgpwngsadaagigkyayayidikeda 1000
OY 129 GCINIVRDGTDKLIDSDLRVAFGDFTRTVSVIAGNSAVYDSRADAFRAFGVALAEAH 188
DB 1001 nkigflfyn-----kgsqgqgtgdmtdmlik----- 1025
OY 189 WVDKRTLLMPGGODKPIVRLYS-----HSKRVADGEGKFTDRYLKLTPTTVSOQVSMR 243
DB 1026 --qynqlfvekedkvynpypcvplavlsgevlsd-----klisllftrlegld 1073
OY 244 FPHLSVYAFKLPDNAV--DELLOGETVAIAAE--DGILISATQV-QTACV----- 291
DB 1074 leelkegleikdvgnsdsvfclvtelegctvhvgefdiekpfsvtyldrtlsvksqwk 1133
OY 292 -LDDIAAEAEELSTGAQL-ADGVTFRVWAPTAQOVVVVYSA--DKVIGSHPTKDS 347
DB 1134 lldemyaydgk--lgaehedgatalikwspkadnsvvlydkvdqnevdliemvkgd 1190
OY 348 ASGAMSWO-----GGSDDKGAFAFYARMVYHPSKRVQYEVET--DPVAHSLSST-NS 396
DB 1191 -rgvsvvkltkdntgldskikyynhyelc-----hgdvtnlaldpyskmaawm 1239
```

```
OY 397 E-----YQOVVLDNSALKPDGMDNLMPHAOKTKADIAKMTIHESHIRDSAMDQVTP 450
DB 1240 eagdkvgkaalvdi--gsigpe-lidyadipgfekecdtl-----yehvtrdfcs-dpnlg 1291
OY 451 AELRGYVLAITLTDGDSNMVQHKLTLSAGVTHVELLPVFPDLATVYNESDKVADIQPFSL 510
DB 1292 edlkafgfrf--asiveklidygelqlipvmysyfsnef-----eegexrm 1339
OY 511 CEVNSAVKSSSEFAGYDSDSTVEEVNLQKSDSDONPOVALNLTVAQTDSYNMGYDPE 570
DB 1340 le-----yasgl-----nywgydph 1356
OY 571 HYTPEGSYATDPEG-TTRIKERTMIQAI-KODLGMNIMDVVYNHTNAAGPTDRTSVL 628
DB 1357 nyfslsgmysenpedpelikeeknlinslhrkdmg--vldvfrnh-----taqvhl 1408
OY 629 DKIVPWTYORLNETTGSVSATCCSDSAPEHNFALKLIDSLAVWTTDKIDGFRFDLNG 688
DB 1409 edlvprnyhfm-dadgltprtsfggrglgtlhemarrvlvdsllhwvdeykvvdgfrfdmg 1467
OY 689 YHPKAQILSAMERIKALNDIYFFGCGW--DSNOSDREIASQIMLKG-TGIGTFSDDL 744
DB 1468 dhdaesiqlafdeaklmpnlmimgewvcltagdegepvqaadqwmqyleavgsisdef 1527
OY 745 RDSVRGSGPFSDGDLRQNGIGSGAGVLPNELASISDDQVRLADLTRIGMAGNLADPV 804
DB 1528 rnelk-----sfgf-----segqprfllt----- 1545
OY 805 MIDKCGAKKSGEI--DYKAPGEGVAD--PTEVNVTVSKHNDOTLMDMSTYKSGEADLA 861
DB 1546 -----9gavavvqglfdnlikaplnfnadqpgdvgyleahdnlllydvlaqslkkdpela 1600
OY 862 -----TVVRMQAVSLATVMLGGGIAFPQGSSEILRSKS----- 894
DB 1601 endlrlkrlr--vgnamvltsqgtaflhagqefgrlqwrarapateapykstymldadg 1657
OY 895 -----FTRDSYDSGDMFRVDYSLQDNNVNYGMPRISDDGSNVEVITRVKEMVATPGE 947
DB 1658 nptfvyryfihdsydsdlinrfdwk-----atdaekypvnnvtrdyta----- 1701
OY 948 AELKQMTAFYQELTELKRSPLFTLCDSAVMKRVDFRNTGSDQAGLLVMTVDDGMKAG 1007
DB 1702 -----qllelrssdsdfrlgsrelvdnsnvmvdapelkeqdlv-----ay 1742
OY 1008 ASLDSRLDGLVVAINAPESTRINEFAGETIQLSAQOTAGENSILAN-----GVQIAAG 1062
DB 1743 rsvslagvayylfvmadlsrllt--lqgdllegvvvvaadeaanvagvaepagfeltaeg 1800
OY 1063 TVTLPAWSVAVLELPOGEAG 1083
DB 1801 -ltleplttvvrv--gegeg 1818
```

Search completed: April 22, 2001, 08:57:42  
Job time: 7283 sec

• • • • •





Db 623 ALLGKDTMSKASSELHAINPGIALYGEPTWGTSAIPDDQLITKGAKOKGVAAYVNDNR 682  
 Oy 746 DSVRGCGPPDSGDALRONOGIGSGAGVLPNELASLDDQVRLADLTRLGAGNLADPVM 805  
 Db 683 NAL-DGNVFDSDSAGGFATGA-----TGLTDAIKNGVEGSINDF-- 719  
 Oy 806 IDKGAARKGSEIDYNGAPRGYAADPFEVNVYSKHDNOTLMDMSTYASQADLATTR 865  
 Db 720 -----TSSPGETINVTSHDNTLTMDKIALSNPNDS-ADRIK 756  
 Oy 866 MQAVSLATVMLGGIAFDQGSSELRKSFTRDSYSGDMFNRVYSLDDNNYNYGMRI 925  
 Db 757 MDLQAQAVVMTSQGVPEFGGSEMLRKXGNDNSTNAGDAVNEPWS----- 803  
 Oy 926 SDDGSNEYITRVKEMVATPGAEELKONTAFYOEITELRKSSPLFTLGDSAVMKRVDR 985  
 Db 804 -----RKAQYPDVENYYSGLIHLRLDHPAFRMTTANINSHLQFL 843  
 Oy 986 NTGSDQOAGLLVMTVDGKMKAGASLDSRLDGLVVAINAPESRTINEFAGETLQLSAIQ 1045  
 Db 844 NSPENTVAEYELTDHVN-----KDKWGNIIYVYNPNKTVATINLPSGK-MAINATSG 893  
 Oy 1046 TAGENSLANGVQIADGTVTLPAMSAVL 1074  
 Db 894 KVGESTLGO-----AEGSVQYFGISMIL 917

RESULT 3  
 US-08-477-630-11  
 ; Sequence 11, Application US/08477630  
 ; Patent No. 5721128  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DEMEER, PHILIPPE  
 ; APPLICANT: AMORY, ANTOINE  
 ; TITLE OF INVENTION: PULLULANASE, MICROORGANISMS WHICH  
 ; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 ; TITLE OF INVENTION: PULLULANASE AND THE USES THEREOF  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOPER GILSON & LIONE, P.C.  
 ; STREET: 2000 K Street, N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/477,630  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: William F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-42  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 429-0625  
 ; TELEFAX: (202) 293-1850  
 ; TEXT: 650 383-5605  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 928 amino acids  
 ; TYPE: amino acids  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-477-630-11

Query Match 14.0%; Score 788.5; DB 1; Length 928;  
 Best Local Similarity 25.4%; Pred. No. 4,7e-56;

Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;  
 Oy 77 AADYSKMLYLN-----NETCDALSAVPA-----DW-NDYST 108  
 Db 15 AGDYQPSLIMMPKRGKGAGYEYFNQPADSFGVAADIPGNISQVGIIVRTDRTKDS- 73  
 Oy 109 TPGSGDKY-----GPVVIPLNKE-----SGCINVIYRQDTRKLDISDLRV 149  
 Db 74 ----ADRYIDLKSGNEVLMVEGNSQIFYNEKDAEDAAKPAVSNAVILDSNOYVLKLSQPL 129  
 Oy 150 AFGD-----FT--DRVSYIAGNSAVYDSR-----ADAFRAAFGVALAAHAYDKN- 193  
 Db 130 TLGEGXSGFTVADDPANODIPVTSVKDASLGODVTAVALAGTFQHLFG-----GSDAPDMH 185  
 Oy 194 -TLMPGGODKPIVRLYLSHSSKVAADGKPF-----TDRLKLPTTVSQOVSMPRP- 245  
 Db 166 STLL-----KVTNNLQFSGDLF--EGNQYKVALNDSR--NPFSPSNNILYPA 234  
 Oy 246 ---HI-----SSYAFKLPDANAVDELLOGET-----VAIAAEDGILLSATQVOTAG 290  
 Db 235 GGAHTFESYIPSTHAYVDITINPNAD--LQVESGKTLDTVTGLGEDPDVSHLSIQTDG 292  
 Oy 291 -----VDDA-YAFAEALSYGAOLADGVTFRMAPTAQOVDYVYV--SADKRYI 338  
 Db 293 YQAKOVIPRNVINSSQYVYSGDDL--GNTYQKAPTFKMAPTSTQVNVLLYDSATGVT 350  
 Oy 339 GSHPMTRDSASGAMSGGSDLKGAFAFYRAMTVHPQSRKVEQYEVTPYASLSTNSET 398  
 Db 351 KIVPMFTA-SGHQVWEATYQNLNENNYTMYEYGOOSTRAV-----DPYATAIAPNGTR 403  
 Oy 399 SOVVDLNDLSALKPDGMDNLTMPHAKOTYADLAKMTIHESHIRDSAMDQTVPAELRGYVL 458  
 Db 404 GMIVDL--AKTDPAGMNS-----DKHITPKNIEDIYEADVDVDFSI--DPSGKMKKKGXKL 456  
 Oy 459 AL-----TAGDSMNOHKLKLSAGVTHVELLPYFDLAYNESDKVADYIQOFSRLCEYN 514  
 Db 457 ALTEKGTGPDVWKYGTGIDSLKQLGITHVQLMPFASNSVDE--TDPTQD----- 503  
 Oy 515 SAVKSSEFAGYCDSGSTVEEVLNQLKQSDSDQDNPOVALNTLVADTDSYMGQDPFHVTY 574  
 Db 504 -----NMQYDRANDYV 514  
 Oy 575 PEGSYATPDEGTTRIKEPRTMIOAI--KODLGNVIMDVYVYHTNNAAGPTDRTSVLDKIYV 633  
 Db 515 PEGQYATNANGNARIKEFEKENVLSLREHIGVN--MDVYVNTFATQISD-----FDKIYV 568  
 Oy 634 WYQRLNETGTSVESATCCSDSAPHEHMA-----KLIAQSLAWYTDYKIDGRFDLM 687  
 Db 569 EYTYR-----TMMQVITPTDQVLEMKLAERPNOVKFTIIDLKTYWVNEYRHIDGRFDLM 622  
 Oy 688 GYHPRQAQLISAMERIKALNPDIYFFGEGWDSNQS--DREIISOJNLKGTGIGTFSRRLR 745  
 Db 623 ALLGKDTMSKASSELHAINPGIALYGEPTWGTSAIPDDQLITKGAKOKGVAAYVNDNR 682  
 Oy 746 DSVRGCGPPDSGDALRONOGIGSGAGVLPNELASLDDQVRLADLTRLGAGNLADPVM 805  
 Db 683 NAL-DGNVFDSDSAGGFATGA-----TGLTDAIKNGVEGSINDF-- 719  
 Oy 806 IDKGAARKGSEIDYNGAPRGYAADPFEVNVYSKHDNOTLMDMSTYASQADLATTR 865  
 Db 720 -----TSSPGETINVTSHDNTLTMDKIALSNPNDS-ADRIK 756  
 Oy 866 MQAVSLATVMLGGIAFDQGSSELRKSFTRDSYSGDMFNRVYSLDDNNYNYGMRI 925  
 Db 757 MDLQAQAVVMTSQGVPEFGGSEMLRKXGNDNSTNAGDAVNEPWS----- 803  
 Oy 926 SDDGSNEYITRVKEMVATPGAEELKONTAFYOEITELRKSSPLFTLGDSAVMKRVDR 985  
 Db 804 -----RKAQYPDVENYYSGLIHLRLDHPAFRMTTANINSHLQFL 843  
 Oy 986 NTGSDQOAGLLVMTVDGKMKAGASLDSRLDGLVVAINAPESRTINEFAGETLQLSAIQ 1045  
 Db 844 NSPENTVAEYELTDHVN-----KDKWGNIIYVYNPNKTVATINLPSGK-MAINATSG 893

QY 1046 TAGENSLANGVOIAADGTVTLPAMSVAVL 1074  
 DB 894 KVESTLGO-----AEGSVOPGISMIL 917

RESULT 4

US-08-472-293-11  
 ; Sequence 11, Application US/08472293  
 ; Patent No. 5731174  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DEMEER, PHILIPPE  
 ; APPLICANT: AMORY, ANTOINE  
 ; TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
 ; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 ; TITLE OF INVENTION: PULULANASE AND THE USES THEREOF  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 ; STREET: 2000 K Street, N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/472,293  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: William F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-44  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 429-0625  
 ; TELEFAX: (202) 293-1850  
 ; TELE: 650 383-5605  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 928 amino acids  
 ; TYPE: amino acids  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-472-293-11

Query Match 14.08; Score 788.5; DB 1; Length 928;  
 Best Local Similarity 25.48; Pred. No. 4.7e-58;  
 Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;  
 QY 77 AADYSKNTLYAM-----NETCDALSAVA-----DM-NDVST 108  
 DB 15 AGDYQWMSLMMKXKDGCAETDNPADSGAASADIPNPSOVCITVTQDMTKDVS- 73  
 QY 109 TPGSGDKY-----GPYVVIPLNKE-----SCGIVIVYDGTDLIDSDLRV 149  
 DB 74 ----ADRYIDLKSGNEVLEGSQJFYNEKDAEDAKPAVSNAVYLDASNOVLKLSQPL 129  
 QY 150 APGD-----FT--DRIVSVIAGNSAVYDSR-----ADAPFAACGVALAEHWVDKN- 193  
 DB 130 TLGEGSGCFVHDDTANKDIPVTSVKDASLGQDVTAVLACTFOHIFG---GSDMAPDNH 185  
 QY 194 -TLAMPGQDKPIVRLYYSHSKVAADGEGKF-----TDRLYLKLPPTYSQGVSMRFP- 245  
 DB 186 STLL-----KKVTNNLYQSSGLP---ECNYQYKALANDSW--NNPSTPSDNINILVPA 234  
 QY 246 ---HL-----SSYAAPKFLPDNANVDELLOGET-----VAIAAADGILISATVOVATG 230  
 DB 235 GGAHVTFYSYIPSTHNAVYDTINNPNAD--LQVESGVKTDLTVTVLGDPDPSVHTLSIQTDG 292

QY 291 -----VLDDA-YAAEAALSYGAOLADGCVTERVAPTAQOYDVVVY-SADKKVI 338  
 DB 293 YQAKQVLPBVLNLSQYYSDDL--GNTYQAKFTFKVAPSTOVNVLVYDASATGSYT 350  
 QY 339 GSHPTSDASGANSWOGSDLKCAFTRYANTYVHPQSRKVEQYEVDPYAHSLSTSEY 398  
 DB 351 KIYPMTA--SGHVEATVONLNNMYAYEVTGGSTRVAV-----DPYATAIAPGTR 403  
 QY 399 SOVVDLNDLSALKPDGMDNLTPHAKTKADLAKMTIHESHRLDLSANDQVTPALRCKYL 458  
 DB 404 GMAYDL--ATDPRAGMS-----DKHTPKNIEDVITEMDRDST--DPSGKMKKGYL 456  
 QY 459 AL-----TAGDSNMVQHLKTLASGVTHVELLPVEDLATVNEFSKVDADIQOQFSLCEVN 514  
 DB 457 ALTEKGTGKPDNVKVTGDSLKQIGITHVOLMPVAFANSVDE--TDPQD----- 503  
 QY 515 SAVASFEAGYCDSGSTVEEVLNOLKQSDODNPQVALTLVAQDTSYMGYDPFHYTV 574  
 DB 504 -----NMGYDPRRYDV 514  
 QY 575 PEGSATDPEGTTRIKETPTMIOAI--KODLGMVIMDVYNTHTAAGPTDRTSVLDKIVP 633  
 DB 515 PEGYATNANGNARIKEFEKENVLSLREHIGVN--MDVYNTHTFATQISD---FDRIVP 568  
 QY 634 WYQRLNETTGSVESATCCSDSAPENHMA-----KLADSLAVWTTDYKIDGFRDLM 687  
 DB 569 EYVYR-----TMQVILIPDOVLEMKLAEPRVQFFIIDLKYVNEVHIIDGFRDLM 622  
 QY 688 GYHKAQILSAMERIKALNDIYFEGGMSNOS--DRFELASINLKIGTIGFSRLR 745  
 DB 623 ALLGKDTMSKRASELHAIINGIALYGEPTGTSLPDDLLTKGAQKGVAVFNDNLR 682  
 QY 746 DSVKCGPFDGDLARONOGISGAGVLPNELASLSDQVRHLADLRLMAGNLADFPVM 805  
 DB 683 NAL--DGNVPFS-----SAGPATCA-----TGLDLAKNEGSSINDF-- 719  
 QY 806 IDKGAKKKSEIDYNGAPGAYAADPVEVYVYSKHDNQLMDISYKASQOEDLATRV 865  
 DB 720 -----TSSPETINYVSHONYTLMDIKALSNPNDSR--ADRIK 756  
 QY 866 MOAVSLATVMLGGIAPDOGSSELRKSFTRSDYSGDMFNRYDLQNNVNGPRI 925  
 DB 757 MDELAQAVVMTSGCVPMQGEELKRAKGNDSYNKGDVNEPDS----- 803  
 QY 926 SDDGSNEYITRVKEVATPGEAELEKMTAFYQELTELKRSPLFTLGDGSAVAKRVDFR 985  
 DB 804 -----RKQYDPVFNYSGLHLRLDHPAFRTTANETNSHQFL 843  
 QY 986 NTGSDQAGLLVTVDDGKAGASLDSRLDGLVAINAADESRTLNEFAGETLQLSAIQ 1045  
 DB 844 NSPENTVAYELTQHVN-----KDKMGNIIIVYVNPNTVATINLPSSK-WAINATSG 893  
 QY 1046 TAGENSLANGVOIAADGTVTLPAMSVAVL 1074  
 DB 894 KVESTLGO-----AEGSVOPGISMIL 917

RESULT 5

US-08-474-545-11  
 ; Sequence 11, Application US/08474545  
 ; Patent No. 5736375  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DEMEER, PHILIPPE  
 ; APPLICANT: AMORY, ANTOINE  
 ; TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
 ; TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 ; STREET: 2000 K Street, N.W., Suite 200  
 ; CITY: Washington

```

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,545
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wilhem F. Gadiano, Esq.
REGISTRATION NUMBER: 37,136
REFERENCE/DOCKET NUMBER: 4121-43
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 429-0625
TELEFAX: (202) 293-1850
TELEX: 650 383-5605
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-545-11

```

```

Query Match 14.08: Score 788.5; DB 1: Length 928;
Best Local Similarity: 25.48; Pred. No. 4.7e-58;
Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;

```

```

QY 77 AADYSKMLYLN-----NETCDALSAFVA-----DW-NDYST 108
DB 15 AGDVPMSLMMKPKGCGAEEDFNQPADSPFCASADIPGNSOVGIIVRTDMKRDVS- 73
QY 109 TPSSGSKY-----GPVWVPLNKE-----SGCINIVRDTGKLIDSDLRV 149
DB 74 ----ADRYIDLSKGNVWLVNCSQIFYNKEDADAKPAVSNVLAIDSNQVLVLSQPL 129
QY 150 AFGD-----FT--DRVSVIAGNSAVYDSR-----ADAFRAFGVALAAHWADKN- 193
DB 130 TLGEXSGFTVHDDTANNDIPTVSVDASLGQDVTAVLAGTFQHLFG-----GSDNAPDHN 185
QY 194 -TLMPGGQDKPIVLYLSHSSKVADEGKF-----TDRTYKLPTTVSOQVSRFP- 245
DB 186 STL-----KKVTNNLYQFSGDL-----EGNTQYVALNDSH--NNPSTPSNINILTYVA 234
QY 246 --HL-----SSYAFKLPDNAVYDELLOGET-----VAIAAEDGILISATOVOTAG 290
DB 235 GGAHYTFSTPSTHAYVTINPNAD--LQVESGVKTDLVITLGEDPVSHTLSIQDTG 292
QY 291 -----VLDDA-YAIAEALSYGAQIADGVTFRVMAPTAOQVNDVYV-SADKKYI 338
DB 293 YQAKOVIPRNVLNSOYVYSGDDL--GNTYQKATTFKVMAPTSTOVNVLVDATGVTY 350
QY 339 GSHPTIRDSASGAWMGSGDLKGAIFYRYAMTVYHPSKRYEYVETDPAVSLSTNSEY 398
DB 351 KIVPMTA-SGHCWEMATYVNONLEMYVYEVYGGOSTRAV-----DPYATAIAPNCTR 403
QY 399 SOYVLDLNSALKPDGMDLTMFPAOKTYADLAKMTIHSHIRDLSAMDOTVPAELRGKYL 458
DB 404 GMIVDL--AKTDPACMNS-----DKHITPKNIEDVIEYEDVDFSI-DPNSGMKMKKGXKL 456
QY 459 AL-----TAGDSMMVQHLKTLASGVTHYELLFVPLDLYVNESSDKVADIDQPFSLCEVN 514
DB 457 ALTEGTGCPDVKVTKGIDSLKQLGITHVQLMVFVAFSNSVDE-TDPTQD-----503
QY 515 SAVKSEFAGYCDSSGTYEEVNLQKQSDSDNPQVQALNTLVAGTDSYNMGDPFHYTV 574
DB 504 -----NMGIDPRNDIV 514

```

```

QY 575 PEGSYATDPEGTTRIKERTMTIOAI-KQDLGMNVIMDVVYVYHNTNAGPTDRTSVLDKIVP 633
DB 515 PEGQYATNANGNARIKEFEKREVLSLHREHIGVN--MDVYVNTFATQISD-----EDKIYP 568
QY 634 WYQRLNETTGVSESATCCSDSAPFHRMFA-----KLIAIDLAWTTDYKIKIDGFFDL 687
DB 569 EYVYR-----FMQVYIIPTDVLEKRLAERPVOYKFIIDSLKVMVNYHIDGFFDL 622
QY 688 GYHFAOILSABERIKALNPDIYFPEGWDSNQ--DREIASQJNLKGTGIGFSDRLR 745
DB 632 ALLGDTMSKAASELHAIIPGIALYGEPTGTSALPDQDLTKGAKGMGVAVFNDR 682
QY 746 DSVKGGPFDSCDALRQNGIGSGAGVLPNELASLDDQVRHLADLTRIGMAGNLADFPV 805
DB 683 NAL-DGNYEDS-----SAQGFATGA-----TGLTAAIKGVGSINDF-- 719
QY 806 IKGDAKKGSELDYNGAAGGYAADPTEVYVYVSKHDOTLMDMISTYKASQADLATVR 865
DB 720 -----TSSPGETINVTSHDNTYTLMDKIALSNPNDSF-ADRIK 756
QY 866 MQAVSLATVYMLGOGIAFDQSGSELLRSKSTFSDSDSGDFRVDYSLQDNNYNGMPRI 925
DB 757 MDLQAGVYMTSQVFPFGGGEMLRXKGNDSINAGVAVFEDMS-----803
QY 926 SDDGSNVEYITRVKEMVATPGAEELKQMTAFYQELTELKSSPLFTLGDSGAVMKRVDR 985
DB 804 -----RAQYPDVFNYSGLIHLRLDHPAFRMTTANFINSILQPL 843
QY 986 NTGSDQAGLLVYVDDGKAKASLDSRLDGLVAINAAPESRTLNEFGETLQLSAIOQ 1045
DB 844 NSPEPTVAELVDHVN-----KDKWGNITIVYVNPRTVATINLPBGK-WAINATSG 893
QY 1046 TAGENSLANGVOIADGVTLPAMSAVL 1074
DB 894 KYGESTLQ-----AEGSYQVPGISMIL 917

```

```

RESULT 6
US-08-478-341-11
Sequence 11, Application US/08478341
Patent No. 5617498
GENERAL INFORMATION:
APPLICANT: DEMERY, PHILIPPE
APPLICANT: AMORY, ANTOINE
TITLE OF INVENTION: PULLULANASE, MICROORGANISMS WHICH
PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS
TITLE OF INVENTION: PULLULANASE AND THE USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: WILLIAM BRINKS HOFER GILSON & LIGNE, P.C.
STREET: 2000 K Street, N.W., Suite 200
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,341
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wilhem F. Gadiano, Esq.
REGISTRATION NUMBER: 37,136
REFERENCE/DOCKET NUMBER: 4121-45
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 429-0625
TELEFAX: (202) 293-1850
TELEX: 650 383-5605
INFORMATION FOR SEQ ID NO: 11:

```

SEQUENCE CHARACTERISTICS:  
 LENGTH: 928 amino acids  
 TYPE: amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-478-341-11

Query Match 14.0%; Score 788.5; DB 2; Length 928;  
 Host Local Similarity 25.4%; Pred. No. 4.7e-58;  
 Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;

77 AADYSSKNLYLMN-----NETCDALSAIPA-----DH-NDYST 108  
 15 AGDQPMWLMMPKDGCAEYDFNQPADSPGAVASADIPGNPSCVGIIVPTQDWTXKVS- 73  
 109 TPSCSDKY-----GPYVWIPLNKE-----SGCINIVRDTGDKLIDSLRV 149  
 74 ----ADRTIDLSKGNENWLVGNSQITFNEKDAEDAKPAVSNAYLDAASQVLKLSQPL 129  
 150 AFCD-----FT--DRTVSVIAGNSAVYDSR-----ADAFRAFGVALAEAHWVDKN- 193  
 130 TLGCGXSGFTVHDPTANKDIPVTSVKDASLGQDVTAVLAGTFOHIFG---GSDMAPDNH 185  
 194 -TLLMPGQOKPIVRLYXSHSSKYADGEGK-----TDRYKLITPTTVSQOVSAREP 245  
 186 STLL-----KKYNNLYOFSGLP---EGNYQYKVALNDSW--NNSPSPDNINLTVPA 234  
 246 ----HL-----SSYAAFKLPDANAVDELLOGET-----VAIAAEFGILISATOVOTAG 290  
 235 GGAHVTSYIPSTHAYVDITINNPRAD--LQVESGVKTDLYTLGEPDPDVSHTLSITDGG 292  
 291 -----VLDDA-YAAEAALSYGAOLADGCVTFRWAPTAQOVDVYVY-SADKKVI 338  
 293 YQAKQVIRPVNLSSQYVGGDDL--GNFTYQKATFEKVAPSTQVNVLLYDSATGCVT 350  
 339 GSHPMTDSASGMSWGSDLGKAEFRYAMTYVHPQSRVEQVEQVDDPAHLSSTSEY 398  
 351 KIVPMIA-SGHGWAEATVNOLEMMYMYEVTGCGSTRFAV-----DPYATATAAPGTR 403  
 399 SOYVDLNDLSALKPDGMNLTMPHAKTKADLAKMTIHESHIKLSAMDQVYPAELRGKYL 458  
 404 GMATDL--AKTDPAGMS-----DKHITPKNIEDEVIEMOVRDSI--DPSNGMKNKKYL 456  
 459 AL-----TAGDSNNVOHLKTLISASGVTHVELLPVFDLATAVNEFSDKVADICQPSRLCEVN 514  
 457 ALTEKGTGKPDNVKTCIDLSLKQLGITHVOLMPVAFASNSVDE-TDPTOD----- 503  
 515 SAVKSSFEAGYCGSGTVEEVLNQLKSDSDQDNPOVALMTVLVAGTOSYMMGYPRHYTV 574  
 504 -----NWGYDPRNYDV 514  
 575 PEGSYATDPPECTTRIKERFTMTQAI-KODLGMMVIMDVVYHNTNAAGPTDRTSVLDKIVP 633  
 515 PEGQATNANGMNAIKERFEKENVLSLHREHIGVN--MDVYVNHFTATQISD-----FDKIVP 568  
 634 WYTORLNETTGSVESATCCSDAPBEHMFV-----KLADSLAVMTTDTYKIDGFRFDLM 687  
 569 EYVYR-----TMMQVILIPDOYLEMKLXAEPRVQKFIIDSLKYVWNEVHIIDFRRDLM 622  
 688 GYIPKAOILSAMERIKALNDIYFEGGMSDNGS--DRFEIASQINKKGIIGFGRRLR 745  
 623 ALLCKDTMSKAASELHAINICGALYGERPWGCTGCTALPDQDLTKGAOKGKQVAAVFNINLR 682  
 746 DSVRCGCPFDSCDALRQNOGICSGACVLPNDELASLSDQVRLHLDLTRLGAGNMLADPVM 805  
 683 NAL-DGNVFDG-----SAQGFATGA-----TGLDLAKNGVEGSIINF-- 719  
 806 IDKDGAKKSGSEIDYNGAPCGYAADPTEVYVYKSHNQTLMKTI SYKASQADLALRV 865  
 720 -----TSSPGETINVTSHDNVTPLMDKIALSNPDSE-ADRIK 756  
 866 MQAVSLATVNLGGIAPDQGGSELRKSKFT RDSYDGDWENRVDYSLGDNNTVVGAPRI 925

DB 757 MDELAQAVVMTSOGVPPMOGGEEMLRKKGNDNSYNAGDADVNEEDMS----- 803  
 OY 926 SDDGSNEVITRVKENVATQGEAELKQMTAFYQBELTELKSSPLFTLGDGSAVKKRVDPR 985  
 DB 804 -----RKAQYPOVFNTYSLIHRLDHPAFRMTTANETINSILOFL 843  
 OY 986 NTGSDQAGLLVMTVDDGMRKAGASLSDRLDGLVYVAINAPESRTLANEFAGETLOLSAIOQ 1045  
 DB 844 NSPENTVAELTDHVN-----KDKWGNIIYVVPNNTVATINLPSCG-MAINATSG 893  
 OY 1046 TAGENSLANGVQIADGTVTLPMSAVVL 1074  
 DB 894 KVESTILGQ-----AEGSVQPGISMML 917

RESULT 7  
 US-08-996-733-11  
 : Sequence 11, Application US/08996733  
 : Patent No. 6074854  
 : GENERAL INFORMATION:  
 : APPLICANT: DEMEER, PHILIPPE  
 : APPLICANT: AMORY, ANTOINE  
 : TITLE OF INVENTION: PULULANASE, MICROORGANISMS WHICH  
 : TITLE OF INVENTION: PRODUCE IT, PROCESSES FOR THE PREPARATION OF THIS  
 : NUMBER OF SEQUENCES: 15  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genecor International, Inc.  
 : STREET: 925 Page Mill Road  
 : CITY: Palo Alto,  
 : STATE: CA  
 : COUNTRY: U.S.A.  
 : ZIP: 94304  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/996,733  
 : FILING DATE: 23-DEC-1997  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/472,293  
 : FILING DATE: 07-JUN-1995  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/174,893  
 : FILING DATE: 28-DEC-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09301278  
 : FILING DATE: 19-NOV-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09300744  
 : FILING DATE: 15-JUL-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: BE 09201156  
 : FILING DATE: 28-DEC-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Debra J. Glaister, Esq.  
 : REGISTRATION NUMBER: 33,888  
 : REFERENCE/DOCKET NUMBER: GC446C1-US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (650) 846-7620  
 : TELEFAX: (650) 845-6504  
 : INFORMATION FOR SEQ ID NO: 11:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 928 amino acids  
 : TYPE: amino acids  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-996-733-11



Query Match 14.0% Score 788.5; DB 3; Length 928;  
 Best Local Similarity 25.4%; Pred. No. 4.7e-58;  
 Matches 282; Conservative 151; Mismatches 359; Indels 317; Gaps 46;

77 AADSSKRLYLWN-----NETCALSAIPA-----DW-NDVST 108  
 15 AGDIQPMSLMMMPKDGGAETDFNOPADSFQAVASADIPGNSQYGIYRIODMTRKDS- 73  
 109 TPSSGSKY-----GRVWYPLNKE-----SCCINVIVRDGTDLIDSDLRV 149  
 74 -----ADRIIDLSKNGEVLVEGNSQIFYNKEDAEADAKPAVSNAYLIDANSOVLKLSQPL 129  
 150 AFGD-----FT--DRTVSIYAGNSAVYDSR-----ADAFRAAFGVALAAHAYDKN- 193  
 130 TLGEGXSGFTVDDTPANKDIPVTSVKDASLGODVTAVLAFQHIIFG-----GSDMAPDH 185  
 194 -TLMPGQDPRIVRLVYSHSSKVAADGKGF-----TDRLKLTPTTVSQVNSRPF- 245  
 186 STL-----KVTNNLTFSGDLP---EGNYQYVALNDSM--NNPSTPSONINLTYP 234  
 246 ---HL-----SSYAFLKLPDANVDELLOGET-----VAIAAEDGILLISATQVOTAG 290  
 235 GSAHTFTSYIPSTHAYDITINNPNAD--LQVESGKTDLVITLGEDPVSHTLSIQFDG 292  
 291 -----VLDA-VIAEAALSTGAOLADGVTFRVMAFTAOQVDVYV-SADKKVI 338  
 293 YQAKOVIERNVLSNQYYSDDL--GNTYQKATTFKWAFTQOVNLLYSATGASVT 350  
 339 GSHPTRDSASAWMOGSDLKGAIFYRAMTVYHPOSKVROYEYTDYASLSLNSX 398  
 351 KIVPMTA-SGHGWATYVNONLENNYVMEVGGSTRVAV-----DYATRALANGR 403  
 399 SOVVLNDSALKPDEMDNLTPMHAQTKADAKMTIHESHIDLSAMDQVPAELGKYL 458  
 404 GMIVYL--AKTRPACMNS-----DKHITPKNIDEVITYENDVADFSI--DPSNCKMKNGKYL 456  
 459 AI-----TAGDSMVOHLKTLASGVTHVELLPVFDLAVNEFSKADVIQOPPSRLCEVN 514  
 457 ALTEKGTGPDVVKIGISIKOLGITHVQLMFVAFNSVDE--TDPTQD----- 503  
 515 SAVKSEFAGYCDSSGYEEVLNOLKQSDSDNPQVALNTLVAQTDVSNMGYDPHYTV 574  
 504 -----NMGDPENYDV 514  
 575 PEGSYATPEGTTRIKKEFTMTQAI-KODLGNVIMDVYVNHNTNAGPYRDTSVLDKIYV 633  
 515 PEGOVATNANGNARIKEFEVLSLHREHIGVN--MDVYVNHNTFAQISD---FDKIYV 568  
 634 WYQQLNETTGSVESATCCSDSAPHRMFA-----KLIADSLAVTTTYKIDGFFEDLM 687  
 569 EYYR-----TMQVLIPTDQVLEKMLKAEPRMYOKFIIDSLKVVNNEYHIDGFFEDLM 622  
 688 GVHPRKQILSAMERIKALNPDIYFCEGMDSNOS--DREIASQILNKCTGIGTFSDRLR 745  
 623 ALLGDTMKSASSELHAIIPGIALYGEPTGTSALPDQDLTKGQKMGVAVFMDNIR 682  
 746 DSVRGSGPDSGDLARONOGISGAGVLPNELASLSDDOVRHLADITRLGCMGNLADPYM 805  
 683 NAL-DGNVDS-----SAQGFATGA-----TGTLAIKNGVESINDF-- 719  
 806 IDKGAAKKGSLEIDNGAFGYAADPTEVYVYVSKHDNOTLMDMTSYKASQEDLARTVR 865  
 720 -----TSSPGETINVTSHDNTTLMDKIALSNPNSE-ADRIK 756  
 866 MOAVSLATVMLGOGIAPDOGSELRSKSFTRDSYDSGWMFRNVYSIOLDNNYVNGMPRI 925  
 757 MDELQAVVMTSQVPTFMGGEEMLRXKGGNDNSTINAGAVNEFDMS----- 803  
 936 SDDGSNVEYTRFKEMVATPGAEAILKOMTAFYQELTELKSSPLFTLGDGSAVMKRVDR 985  
 804 -----RKAQYPDVFNYSGLHLHLDPAFKMTTANELINSLIQL 843  
 986 NIGSDQOAGLLVMTVDDGKAKASLDSRLDGLVVAINAAPESRTLNEFAGETLLQLSAIQ 1045

844 NSPENTVAELDHVN-----KDKMGNITVYVNPKNVATINLPSGK-WAINATSG 893  
 1046 TAGENSLANGVOLAADGVTVLPAMSAVL 1074  
 894 KVGESTLGO-----AEGSYQVPGISMAIL 917

RESULT 8  
 US-08-961-083-72  
 ; Sequence 72, Application US/08961083  
 ; Patent No. 6159469  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Choi et. al.  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
 ; NUMBER OF SEQUENCES: 452  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,083  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 72:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 618 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-961-083-72

Query Match 8.0% Score 449.5; DB 4; Length 618;  
 Best Local Similarity 25.2%; Pred. No. 1.5e-29;  
 Matches 165; Conservative 86; Mismatches 208; Indels 197; Gaps 25;

478 GVTHVELLPVFDLAVNEFSKADVIQOPPSRLCEVNSAVKSEFAGYCDSSGYEEVLN 537  
 2 GVTHIQLLPVLISYFVNELKNH----- 23  
 538 QLKQSDSDNPQVALNTLVAQTDVSNMGYDPHYTVVPGSGYATDPEG--TRIKKEFTMT 596  
 24 -----ERLSDYASSNSNMYNGVDPONYSFLTGWSSDPKNEPRIAEFRKNI 70  
 597 QAIKODLGNVIMDVYVYVNHNTNAGPYRDTSVLDKIYPMYQRLNETTGSVESATCCSDSA 656  
 71 NEIHK-RGMGAILIDVYVNH-----TAKVDLFEDELPEPNYHFN-DADGTPRTSFFGGRLG 122  
 657 PEHRMFAKLIADSLAVTTDYKIDGFFEDLMGVHPRKQILSAMERIKALNPDIYFCEGM 716  
 123 TTHNMKRLIIDSIKYLVDTYKVDGFRFDMGHDAASTEAKAKARALNPNIIMLGECH 182  
 717 DSNQSDR-----FEIASQILNKGTG-IGTFSDRLRDSYRGGGPDSGDLARONOGISGAGV 772

```

Db 183 RYVAGENNPTKADODMMKHTDFAVFSDDIRNNLSGVPNGCOPAF----- 230
Oy 773 LPNELASLSDQVYRHLADLTRLGMAGNLADFEVVIDRKDGAKKSEIDYNGAPGYAAD-P 831
Db 231 -----ITGCKRDVNTIFKNLIAQ-----PINFADSP 257
Oy 832 TEVYNTVSKHNDOTLMDMISY-----KASQFADLATRVKQAVSLATVLMGQIAFPDQ 884
Db 258 GGVIOYIAAHNDLTLFDLIAQSIKDKPSKAENYAEIHRRLGNL--MVLTAQGPFIH 314
Oy 885 QGSE-----LHRSKS-----FTRDSVSDSDMNRND 910
Db 315 SCGEYRTQOFNDPAVKTPVAEDKVPNKSHLLDKDGNPFDPYFIHDSVSDSAAVANKFD 374
Oy 911 YSIQDNNYVNGMPRISDDSGSYEVITRVKEMVATPGEAEIKOMT-AF-YOELTEKRKSSP 968
Db 375 WTK-----ATDKAYP--ENVKSROYMKGLIALRQSTDAFLRLSLQIDIDRVH 420
Oy 969 LFTLGDSAVMKRVDFRNTGSDQOAGILVMTVDDCKKAGASLDSRLDGLVVAINAAPESR 1028
Db 421 LTVPGONGVEK-----EDVYIGYQI-TAPNG-----DIYAVFVNADEKAR 460
Oy 1029 TLN---EFAG-ETLOLSAIOQTACENSILAN--GVQIAADQTVTLRPAVSVALELPQ 1078
Db 461 EFMUGTAFAHLRAEVLADENQAGPYGIANPKGLEWTEKG-LKLMALATVTLRVSQ 515

RESULT 9
US-08-476-519-2
: Sequence 2, Application US/08476519
: Patent No. 5750876
: GENERAL INFORMATION:
: APPLICANT: Barry, Gerard F.
: APPLICANT: Kishore, Ganesh M.
: APPLICANT: Krohn, Bradley M.
: TITLE OF INVENTION: No. 5750876el Isoamyrase Gene, Compositions
: TITLE OF INVENTION: Containing It and Methods of Using Isoamyrases
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Grace L. Bonner, Monsanto Company, B44F
: STREET: 700 Chesterfield Parkway No. 5750876ch
: CITY: St. Louis
: STATE: Missouri
: COUNTRY: USA
: ZIP: 63198
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/476,519
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/281902
: FILING DATE: 28-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Bonner, Grace L.
: REGISTRATION NUMBER: 32,963
: REFERENCE/DOCKET NUMBER: 38-21(13577)A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314)537-7286
: TELEFAX: (314)537-6047
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 777 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-476-519-2

```

```

Query Match 6.3%; Score 352.5; DB 1; Length 777;
Best Local Similarity 22.6%; Pred. No. 3,9e-21;
Matches 203; Conservative 102; Mismatches 308; Indels 285; Gaps 37;

Oy 289 AGVLDATAYAAALSYGAAQ--LADGCVTFRRVAPPAQOVVVVYSDAKKVICSHPTTR- 345
Db 24 ACALSPAH-AIDAQOLGARVYAAQANLAFRVYSSRATREVELY---KMPSTSQEVARL 79
Oy 346 ---DSASGAMS-----WQGS-----D 359
Db 80 ALSKDPATQWLSLPTSTIKNTYGITGAVYCYRANGPMPDAAATKTSATGFPYSDVD 139
Oy 360 LKCAFYRYAMTVYHPQSRKVEQYEVTDP-----VAHSLSTNSEVSQVYDLNDSAL 409
Db 140 MAGNRFNPKLLDLPVAREISO---DPNATCATDGTIVATGAHNRKDSGLASKGIAL 195
Oy 410 KPGMDNLTPHAOKTKRADIAKMTIHESHTRDISAMDQYVPAELRGYKALTLTGDSNMVQ 469
Db 196 AADA---TSVGSKPTIRA-LKDEVIEVHVRLTRNDSDVPAAREGTY-----KGAAR 243
Oy 470 HLKTLASGVTHVELLPVFDLATVNEFSKVDADIQQFSLCEVNSAVKSEFAGYCDSG 529
Db 244 KAALALAGTVTAVERLPV----- 261
Oy 530 STVEEVLNQLKQSDSD--NPQVALMTLVAQTDSYNMGYDPFHYTVPEGSYATDPEGTTR 588
Db 262 -----QETQNDQNDVDNSTA-----GDNY-NGWYTLVYFAPDRRAYADKSAGP 305
Oy 589 IKEFRRMIOALIKODLGNNVIMDVVYNNHTNAAP---TDRTSVLDKIV-----PWYORL 639
Db 306 TREMKAVVKA-FHDAGIKVYIDVYVYNTGCGPWSGTDGLSYVYNLSLFRGLDNPAYISLS 364
Oy 640 NETTGSVEATCCSDSAPEHMFAKLIADSLAVWTTDYKIDGRFRDLMGHPAQLISAW 699
Db 365 SDVKYPMNDTGVCGNNTNRPIAQNOLVDSLAYWRDAGVDCGRFDL-----ASVIG-- 416
Oy 700 ERIKALNPDIYFFEGEWDNSQSDRFEIASOINLKGTGIGTFSRLRD--SVRCG---GP 753
Db 417 ---NSCHGCFNFDKNDGNAI--IVAEIPLPPAAGAGADLIAEPMAIGGNSYQVGG 471
Oy 754 FDSG-----DALRONGISGAGVLPNELASLSDQVYRHLADLTRLGMAGNLADFEV 804
Db 472 FPAQMAEMNGLYRDLAKKQKLGVEYVTPGTLA-----TR--FAGS----- 511
Oy 805 MIDKGAARKGSEIDYNGAPGYAADPTEVYVYVSKHNDOTLMDISYKASQEAD----- 859
Db 512 -----NDLVGDDGRKPMHSINFYVAHDCTLNDLVAYNKKQNNQPPYRG 555
Oy 860 -----LATRVKQAVSLATVLMGQIAFDQOGEILRSKSFTRDSYDS 902
Db 556 PSDGGEHNLMSWOGGIVAEORKAARTGLALLMSAGVPMTGDEALFRQ----- 606
Oy 903 GDMFNRVDSLDONNYYVNGMPRISDDGSNYEVITRVKEMVATPGEAEIKOMTAFYQELTE 962
Db 607 ---FG-----NNNTYTNL-----DSANMLYWSR--SALEADHETTYTKRLAFRAKHPA 649
Oy 963 LRKSSPLFTLGDSAVMKRVDFRNTGSDQOAGILVMTVDDCKKAGAAA---SLDSRLDGLV 1018
Db 650 LRANPFASADTGNVNWEOGLRMFKPDGAQ-----ASAVFNGADNNAALAMRIDGSE 700
Oy 1019 VAINAPESRFLNEFAGET-----LQLSAIOQTACENSILANGVOIAADQTVTL 1066
Db 701 FGDSASAIYVAWNGSGAVDFKLPMPTGQKQMYRVETDTATWNGSPNAV--ALPGSETL 756

RESULT 10
PCT-US95-09323-2
: Sequence 2, Application PC/TUS9509323
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Novel Isoamyrase Gene, Compositions
: TITLE OF INVENTION: Containing It and Methods of Using Isoamyrases

```





RESULT 13  
5457037-3  
; Patent No. 5457037  
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;  
; GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO  
; TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE  
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME  
; NUMBER OF SEQUENCES: 7  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/1,797  
; FILING DATE: 08-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 749,621  
; FILING DATE: 19-AUG-1991  
; APPLICATION NUMBER: 224,114  
; FILING DATE: 25-JUL-1988  
; SEQ ID NO:3:  
; LENGTH: 750  
5457037-3

Query Match 5.9%; Score 334; DB 6; Length 750;  
Best Local Similarity 22.5%; Pred. No. 1,4e-19;  
Matches 176; Conservative 86; Mismatches 242; Indels 278; Gaps 32;

QY 304 SYGAQLADGVTFRVMAPTAQOVVVVVSADKKVIGSHPTRDSA-SCAW-----SMOG 356  
DB 10 SYDAQOAN--ITFRVSSQATRIVLVLYSAGYQVESATYTLSPAGSGVMAVTVPVSSIK 67  
QY 357 GSDLKAFRYAMVYVHOSRKVEQYEVTDPAHSLSTNSEYSOV--VDLNDLSALKPDGM 414  
DB 68 AAGTIGAVY-YGYRAMGNW-----PYASNMKGSGOAGFVSDANDRENPN-- 114  
QY 415 DNLTPHAKTKAD-----LAKMT----- 433  
DB 115 KLILDPYAEVSQDPLNSONGNVFASGASYRTTDSGIYAPKGVLYVPSSTGSKPTR 174  
QY 434 -----IHESHRLDSANDQTPAELRGKYLALTAGDSNMVQHLKT--LSASGVTHVELL 485  
DB 175 AOKDDVIEVHVRGFTEDDTSIAPQYRGTYG--AG-----LKASYLASLGTVAVEFL 225  
QY 486 PVFDLATVNEFSKRYADIQOFPESRLCEVNSAVKSEFAGYCDGSGTVEVLNQLKOSDSQ 545  
DB 226 PVQE--TQNDANDVYPN-----SDAN 244  
QY 546 DNPVOALNTLVAQDTSYMGYDPFHVYVPEGSYATDEGTRRIKEFRMTIOAIRKODLGM 605  
DB 245 QN-----YMGYMTENYFSPDRRYAVNKAAGGPTAEFGAMVOAF--HNAGI 287  
QY 606 NVIMDVVYNHTNAAG-----PTDRTSVLDKIVPW-----YQRLNETTGSVESATCC 652  
DB 288 KYVMDVVYNHTAEGETWTSSDPTAT-----IYSWRGLDNTTYVELTSGNOFYFYNTGIG 342  
QY 653 SDSAPEHMFAPKLADSLAVVTTDYKIDGFRFDLMGYHPKAOILSAMERIKALNPDIFYF 712  
DB 343 ANFNNTYVNAOQLIVDSLAVYANTMGVDFRFDLASVLSGNSCLNKAIVYASAPNCNGGIN 402  
QY 713 GEGWDSNOSDFEELASQINLK-----GTGIGTFSSRLR--DSVRGCGPFDG----- 757  
DB 403 FDADSNVAIN-RILRETVRPAAGSGCLDFAEPMAIGNSYQLG--FPGQWSEWNGLF 460  
QY 758 -DALPQNGIGSGAGVLENEIASLSDDOVRHLADLTRLGMAGNLADFPVIMDKGAARKGS 816  
DB 461 RDSLROAO-----NELGSMT-----IYVYODAN----- 483  
QY 817 EIDVNGAPGAYAA--DPEVYVYVSKHDNOTLWDMIS----- 851  
DB 484 --DPSGSSNLTFOSSGSRWSINFLIVHDKMTLKVYSCGANNSSQAMPYPSDGGTSTN 541  
QY 852 YKASQEAADLAT-----RYRMQAVSLATVMLGQIAFDQGSSELLRSKFTTRDSYSGDMF 906  
DB 542 YSMQGSAGTGAIVNQRARARTGMAFEMLSAGTFLMGGGDEYLTLOCNNNAIVMLDSSA 601

QY 907 NKVDYSIQDNNVNGMPRISDGSNYEVTTRYKENVATPGEAEKLQMTAFQOELTELKRS 966  
DB 602 NMLTVYSW-----TTQDSNF-----YTFQAQLLAFKRA 628  
QY 967 SP 968  
DB 629 HP 630

RESULT 14  
5457037-5  
; Patent No. 5457037  
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;  
; GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO  
; TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE  
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME  
; NUMBER OF SEQUENCES: 7  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/1,797  
; FILING DATE: 08-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 749,621  
; FILING DATE: 19-AUG-1991  
; APPLICATION NUMBER: 224,114  
; FILING DATE: 25-JUL-1988  
; SEQ ID NO:5:  
; LENGTH: 751  
5457037-5

Query Match 5.9%; Score 334; DB 6; Length 751;  
Best Local Similarity 22.5%; Pred. No. 1,4e-19;  
Matches 176; Conservative 86; Mismatches 242; Indels 278; Gaps 32;

QY 304 SYGAQLADGVTFRVMAPTAQOVVVVVSADKKVIGSHPTRDSA-SCAW-----SMOG 356  
DB 11 SYDAQOAN--ITFRVSSQATRIVLVLYSAGYQVESATYTLSPAGSGVMAVTVPVSSIK 68  
QY 357 GSDLKAFRYAMVYVHOSRKVEQYEVTDPAHSLSTNSEYSOV--VDLNDLSALKPDGM 414  
DB 69 AAGTIGAVY-YGYRAMGNW-----PYASNMKGSGOAGFVSDANDRENPN-- 115  
QY 415 DNLTPHAKTKAD-----LAKMT----- 433  
DB 116 KLILDPYAEVSQDPLNSONGNVFASGASYRTTDSGIYAPKGVLYVPSSTGSKPTR 175  
QY 434 -----IHESHRLDSANDQTPAELRGKYLALTAGDSNMVQHLKT--LSASGVTHVELL 485  
DB 176 AOKDDVIEVHVRGFTEDDTSIAPQYRGTYG--AG-----LKASYLASLGTVAVEFL 226  
QY 486 PVFDLATVNEFSKRYADIQOFPESRLCEVNSAVKSEFAGYCDGSGTVEVLNQLKOSDSQ 545  
DB 227 PVQE--TQNDANDVYPN-----SDAN 245  
QY 546 DNPVOALNTLVAQDTSYMGYDPFHVYVPEGSYATDEGTRRIKEFRMTIOAIRKODLGM 605  
DB 246 QN-----YMGYMTENYFSPDRRYAVNKAAGGPTAEFGAMVOAF--HNAGI 288  
QY 606 NVIMDVVYNHTNAAG-----PTDRTSVLDKIVPW-----YQRLNETTGSVESATCC 652  
DB 289 KYVMDVVYNHTAEGETWTSSDPTAT-----IYSWRGLDNTTYVELTSGNOFYFYNTGIG 343  
QY 653 SDSAPEHMFAPKLADSLAVVTTDYKIDGFRFDLMGYHPKAOILSAMERIKALNPDIFYF 712  
DB 344 ANFNNTYVNAOQLIVDSLAVYANTMGVDFRFDLASVLSGNSCLNKAIVYASAPNCNGGIN 403  
QY 713 GEGWDSNOSDFEELASQINLK-----GTGIGTFSSRLR--DSVRGCGPFDG----- 757  
DB 404 FDAADSNVAIN-RILRETVRPAAGSGCLDFAEPMAIGNSYQLG--FPGQWSEWNGLF 461  
QY 758 -DALPQNGIGSGAGVLENEIASLSDDOVRHLADLTRLGMAGNLADFPVIMDKGAARKGS 816  
DB 462 RDSLROAO-----NELGSMT-----IYVYODAN----- 484



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 22, 2001, 09:04:38 ; Search time 62.87 Seconds  
(without alignments)  
1192.568 Million cell updates/sec

Title: US-09-262-126c-4  
Perfect score: 5633  
Sequence: 1 MURYNALVLSGLVLLSGC.....AVLELPQGAQAGLPPVSK 1091

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database : PIR-67:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5618.5	99.7	1090	2	alpha-dextrin endo
2	4778.5	84.8	1096	2	alpha-dextrin endo
3	4435.5	78.7	1038	2	alpha-dextrin endo
4	1503	26.7	964	2	pullulanase - spin
5	1432.5	25.4	962	2	starch debranching
6	1424.5	25.3	904	2	probable alpha-dex
7	1405.5	25.0	910	2	alpha-dextran endo
8	1326.5	23.5	986	2	alpha-dextrin endo
9	703.5	12.5	843	2	pullulanase - Ther
10	647.5	11.5	718	2	pullulanase amyl -
11	605.5	10.7	1072	2	alkaline amylopull
12	593.5	10.5	717	2	pullulanase amyl
13	352.5	6.3	716	2	probable glycosyl
14	352.5	6.3	721	2	probable glyx prot
15	344.5	6.1	720	2	glycogen operon pr
16	334	5.9	720	2	glycogen operon pr
17	334	5.9	776	2	isoamylase (EC 3.2
18	332.5	5.9	776	2	isoamylase (EC 3.2
19	328.5	5.8	656	2	glycogen operon pr
20	314.5	5.6	718	2	glycogen operon pr
21	302.5	5.4	664	2	glycosyl hydrolase
22	291	5.2	657	1	glycogen operon pr
23	289.5	5.1	666	2	probable glycosyl
24	280.5	5.0	707	2	glycogen operon pr
25	279.5	5.0	666	2	glycosyl hydrolase
26	266	4.7	789	2	probable isoamylas
27	266	4.7	746	2	glycogen operon pr
28	264	4.7	818	2	probable isoamylas
29	242.5	4.3	659	2	glyx protein (EC 3
			552	2	probable alpha-glu
					S19134

30	233	4.1	598	2	S65770	maltooligosyltreha
31	221.5	3.9	561	2	S73087	alpha-amylase (EC
32	217.5	3.9	596	2	JC4697	trehalase trehaloh
33	206.5	3.7	583	2	G83375	probable glycosyl
34	187	3.3	2059	2	D82671	surface protein XF
35	184.5	3.3	558	2	JC5135	alpha-amylase (EC
36	176	3.1	2334	2	S32920	cell wall-associat
37	170	3.0	2383	2	D64962	probable membrane
38	169.5	3.0	580	2	D70363	1,4-alpha-glucan b
39	169	3.0	630	2	G70763	probable maltoolig
40	169	3.0	600	2	H75516	probable maltoolig
41	169	3.0	1190	2	A82615	surface protein XF
42	167.5	3.0	666	2	B56639	1,4-alpha-glucan b
43	166.5	3.0	731	2	B70770	probable glyx prot
44	165	2.9	3029	2	S76109	hypothetical prote
45	163	2.9	730	1	I64118	1,4-alpha-glucan b

## ALIGNMENTS

RESULT 1  
S11823  
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) precursor - Klebsiella pneumonia  
N:Alternate names: pullulanase  
C:Species: Klebsiella pneumoniae  
C>Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text-change 08-Oct-1999  
C:Accession: S11823; MUID:90205629  
R:Kornacker, M.G.; Pugsley, A.P.  
Mol. Microbiol. 4, 73-85, 1990  
A:Title: Molecular characterization of pulA and its product, pullulanase, a secreted  
A:Reference number: S11823; MUID:90205629  
A:Accession: S11823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1090 <KOR>  
A:Cross-References: EMBL:X52181; NID:943912; PIDN:CAA36431.1; PID:943913  
R:D'Enfert, C.; Pugsley, A.P.  
J. Bacteriol. 171, 3673-3679, 1989  
A:Title: Klebsiella pneumoniae pulS gene encodes an outer membrane lipoprotein regul  
A:Reference number: A32880; MUID:89291709  
A:Accession: A32880  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 944-1090 <DEN>  
A:Cross-References: GB:M29097; NID:9342046; PIDN:AA61976.1; PID:9643577  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 99.7%; Score 5618.5; DB 2; Length 1090;  
Best local Similarity 99.9%; Pred. No. 0;  
Matches 1090; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	1	MURYNALVLSGLVLLSGCNDGSSSSGNDPDPNDVYRLPDVAAGCAVAVEN	60
DB	1	MURYNALVLSGLVLLSGCNDGSSSSGNDPDPNDVYRLPDVAAGCAVAVEN	59
QY	61	QAVYHVDIAGITSSSAADYSSKNLYLWNETCDALAPVADNDVSTPSSGDKGPYV	120
DB	60	QAVYHVDIAGITSSSAADYSSKNLYLWNETCDALAPVADNDVSTPSSGDKGPYV	119
QY	121	VIPLNESGGINVVRDGTDLIDSDLRAFGDFTORTVSVVIGNSAVYDSRDAFRAAF	180
DB	120	VIPLNESGGINVVRDGTDLIDSDLRAFGDFTORTVSVVIGNSAVYDSRDAFRAAF	179
QY	181	GVALLAEHWVWKNTLLMPGQDKPIYRLYSSSKVAADGKFTDRLTLPTTVSQOV	240
DB	180	GVALLAEHWVWKNTLLMPGQDKPIYRLYSSSKVAADGKFTDRLTLPTTVSQOV	239
QY	241	SMFPHLSVYAAFKLPDNAVDELLOGETVAIAAEDGILISATOVOTAGVLDAAEAA	300
DB	240	SMFPHLSVYAAFKLPDNAVDELLOGETVAIAAEDGILISATOVOTAGVLDAAEAA	299

```
Qy 301 EALSTGAGLADGCTFRVMAPTAQOVVVVYSADKKVYIGSHPMTRDSASGAMSGQGSGL 360
    |||
Db 300 EALSTGAGLADGCTFRVMAPTAQOVVVVYSADKKVYIGSHPMTRDSASGAMSGQGSGL 359
Qy 361 KGAFYRATMYHPOSKRKEVEYETDPYALISLSTNSEYSOVVDLNDLSALRKPDGMDNLTMP 420
    |||
Db 360 KGAFYRATMYHPOSKRKEVEYETDPYALISLSTNSEYSOVVDLNDLSALRKPDGMDNLTMP 419
Qy 421 HAAKTKADLAKMTIHESHIRLDSAMDQTPAELRGKYLATLTDGSDNNVOHLKTLTSLASGVT 480
    |||
Db 420 HAAKTKADLAKMTIHESHIRLDSAMDQTPAELRGKYLATLTDGSDNNVOHLKTLTSLASGVT 479
Qy 481 HVELLPVFDLATVNEFSQKADVADIQPFSLCEVNSAVKSSSEFAATCCSGSTVEEVLNOLK 540
    |||
Db 480 HVELLPVFDLATVNEFSQKADVADIQPFSLCEVNSAVKSSSEFAATCCSGSTVEEVLNOLK 539
Qy 541 QSDSQDNPQVQALNTLVAQOTDSYNNMGYDPFHYTEPESYATDPGCTTRIKERFRMIOAIK 600
    |||
Db 540 QSDSQDNPQVQALNTLVAQOTDSYNNMGYDPFHYTEPESYATDPGCTTRIKERFRMIOAIK 599
Qy 601 QDLGMVIMDVYVYHNTNAGPFTDRTSVLDKIVPWYQRLNETTGSVESATCCSDSAREHR 660
    |||
Db 600 QDLGMVIMDVYVYHNTNAGPFTDRTSVLDKIVPWYQRLNETTGSVESATCCSDSAREHR 659
Qy 661 MFAKLIDSLAVWTTDYKIOGFRFDLMGYHRAKQIILSAMERIKALNDIYFPGGCMDSNQ 720
    |||
Db 660 MFAKLIDSLAVWTTDYKIOGFRFDLMGYHRAKQIILSAMERIKALNDIYFPGGCMDSNQ 719
Qy 721 SDREIASQIINLKGTIGTFSRLRDSVRGGGPFPSGDALRONGIGSAGVLPNELASL 780
    |||
Db 720 SDREIASQIINLKGTIGTFSRLRDSVRGGGPFPSGDALRONGIGSAGVLPNELASL 779
Qy 781 SDDQVRHLADLTRLGMAGNLADFYVIMDKGAKKGEIDYNGABGYAADPTEVYVYYSK 840
    |||
Db 780 SDDQVRHLADLTRLGMAGNLADFYVIMDKGAKKGEIDYNGABGYAADPTEVYVYYSK 839
Qy 841 HDNQTLMWDMISYKASOEADLATRYRMQAVSLATYMLGOGIAFDQGSSELSKSFTRDSY 900
    |||
Db 840 HDNQTLMWDMISYKASOEADLATRYRMQAVSLATYMLGOGIAFDQGSSELSKSFTRDSY 899
Qy 901 DSGDMFRVNDYSLQDNNYNNMGPRISDDGSNEYITREKKEKAVATPGAEELKOMAFYQEL 960
    |||
Db 900 DSGDMFRVNDYSLQDNNYNNMGPRISDDGSNEYITREKKEKAVATPGAEELKOMAFYQEL 959
Qy 961 TELKSSPFLTLGDSAVMKRVDFRNTGSDQAGLLVTVDDGKAKAGASLDSRLDGLVVA 1020
    |||
Db 960 TELKSSPFLTLGDSAVMKRVDFRNTGSDQAGLLVTVDDGKAKAGASLDSRLDGLVVA 1019
Qy 1021 INAAPESTLNEFAGETLQLSAIQOTAGENSANGVOIADGVTYTLPAWSVAVLELPOGE 1080
    |||
Db 1020 INAAPESTLNEFAGETLQLSAIQOTAGENSANGVOIADGVTYTLPAWSVAVLELPOGE 1079
Qy 1081 AAGAGLPVSSK 1091
    |||
Db 1080 AAGAGLPVSSK 1090

RESULT 2
alpha-dextrin endo-1,6-alpha-glucosidase (Ec 3.2.1.41) precursor - Klebsiella pneumoniae
N:Alternate names: pullulan 6-glucanohydrolase; pullulanase
C:Species: Klebsiella pneumoniae
C:Date: 30-Jun-1988 #sequence_rev: 30-Jun-1988 #text_change: 06-Dec-1996
C:Accession: A26879; S02472
R:Katsuragi, N.; Takizawa, N.; Murooka, Y.
J. Bacteriol. 169, 2301-2306, 1987
A:Title: Entire nucleotide sequence of the pullulanase gene of Klebsiella aerogenes W70.
A:Reference number: A26879; MUID:87194626
A:Note: K. aerogenes
A:Accession: A26879
A:Molecule type: DNA
A:Residues: 1-1096 <KAT>
A:Experimental source: strain W70
```

```
R:Charalambous, B.M.; Keen, J.N.; McPherson, M.J.
EMBO J. 7, 2903-2909, 1988
A:Title: Collagen-like sequences stabilize homotrimers of a bacterial hydrolase.
A:Reference number: S02472; MUID:99030658
A:Accession: S02472
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-74 <CHA>
A:Note: part of this sequence was confirmed by protein sequencing
C:Genetics:
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-19/Domains: signal sequence, #status predicted <SIG>
F:20-1096/Product: alpha-dextrin endo-1,6-alpha-glucosidase #status predicted <MAT>

Query Match      84.8%; Score 4778.5; DB 2; Length 1096;
Best Local Similarity 84.5%; Pred. No. 9.6e-282;
Matches 934; Conservative 49; Mismatches 77; Indels 45; Gaps 9;

Qy 1 MLRTYRNALVGLSLVLLSGCDNGSSSSSSS-----GNPDTPDNQDVYRLPDVA 49
    |||
Db 1 MLRTYCHALFGLSLVLLSGCDNSSSSSTSGSPGNGNPGTGTGTPDQDVYRLPDVA 60
Qy 50 VPGEAVTAVENQAVIHLVDIAGITSSSADYSSKNLYLNNETCDALSAVADNDVSTT 109
    |||
Db 61 VPGEAVTASARQAVIHLVDIAGITSSPADATKLYLNNETCDALSAVADNDVSTT 120
Qy 110 PGSDKGYPVYVILPLNKESGGINIYRDGTDLKLDSDLRVAFGFTDRTVIVNGSNVY 169
    |||
Db 121 PTGSDKGYPVYVILPLNKESGINIYRDGTINKLIDSC-RKSPDPTDRTVIVNGSNVY 179
Qy 170 DSRADFAFAAGVVALAEAHWDKNTLLMPGGDKPIYELYSHSSKYAADGKFTDRL 229
    |||
Db 180 DSRADFAFAAGVVALAEAHWDKNTLLMPGGENKPIYRLYSHSSKYAADGKFTDRL 239
Qy 230 KLTPTTVSOOVSMKRPPLISYAAFKLPDANVDELQGETVAAIAADGTLI----SATQ 285
    |||
Db 240 KLTPTTVSOOVSMKRPPLISYAAFKLPDANVDELQGGDGIAES-DGILSLSHPGADR 298
Qy 286 VQTGAVLDADAEEAAEALSYGAOLADGVTFRVMAPTAQOVVVVYSADKKVIGSHPMTR 345
    |||
Db 299 RRARRYL-----CRRAEALSYGAOLTDGCVTFRVMAPTAQOVVLYTSADKKVIGSHPMTR 354
Qy 346 DSASGAMSGQSDLKAFYRYATVYHPOSKRKEVEYETDPYALISLSTNSEYSOVVDL 405
    |||
Db 355 DSASGAMSGQSDLKAFYRYATVYHPOSKRKEVEYETDPYALISLSTNSEYSOVVDL 414
Qy 406 DSAKLPQGMNLTMPHAKTKADLAKMTIHESHIRLDSAMDQTPAELRGKYLATLTDGDS 465
    |||
Db 415 DSAKLPQGMNLTMPHAKTKADLAKMTIHESHIRLDSAMDQTPAELRGKYLATLTDGDS 474
Qy 466 NMVOHLKTLASGVYHVELLPVFDLATVNEFSQKADVADIQPFSLCEVNSAVKSSSEFAGY 525
    |||
Db 475 NMVOHLKTLASGVYHVELLPVFDLATVNEFSQKADVADIQPFSLCEVNSAVKSSSEFAGY 534
Qy 526 CDGSGTVEEVLNOLKQSDQDNPOVQALNTLVAQOTDSYNNMGYDPFHYTEPESYATDPEG 585
    |||
Db 535 CDGSGTVEEVLNOLKQSDQDNPOVQALNTLVAQOTDSYNNMGYDPFHYTEPESYATDPEG 594
Qy 586 TTRIKERFTMIOAIKODIGMNVIMDVYVYHNTNAGPFTDRTSVLDKIVPWYQRLNETTGS 645
    |||
Db 595 TTRIKERFTMIOAIKODIGMNVIMDVYVYHNTNAGPFTDRTSVLDKIVPWYQRLNETTGS 654
Qy 646 VESATCCSDSAPAEHRMFAKLADSLAVWTTDYKIDGFRFDLMGYHRAKQIILSAMERIKAL 705
    |||
Db 655 VESATCCSDSAPAEHRMFAKLADSLAVWTTDYKIDGFRFDLMGYHRAKQIILSAMERIKAL 714
Qy 706 NPDIYFEGEGWDSNQSDRFETIASQINLKGTIGTFSRLRDSVRGGGPFPSGDALRONOG 765
    |||
Db 715 NPDIYFEGEGWDSNQSDRFETIASQINLKGTIGTFSRLRDAVAGGPFPSGDALRONOG 774
Qy 766 IGSAGVLPNELASLDDQVYHHLADLTRLGMAGNLADFYVIMDKGAKKGEIDYNGAPG 825
```



[illegible]

RESULT 3  
S38801  
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) precursor - *Klebsiella pneumoniae*  
N|Alternate names: pullulanase  
C|Species: *Klebsiella pneumoniae*  
A|Variety: ATCC 15050; strain K21  
C|Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 08-Oct-1999  
C|Accession: S38801; A25025; S04864  
R|Jansen, B.J.H.; Steyn, A.J.C.; Pretorius, I.S.  
Curr. Genet. 24, 400-407, 1993  
A|Title: Regional sequence homologies in starch-degrading enzymes.  
A|Reference number: S38801; MUID:94130317  
A|Molecule type: DNA  
A|Residues: 1-1038 <JAN>  
A|Cross-references: EMBL:119312  
A|Experimental source: ATCC 15050  
R|Chapou, C.; Raibaud, O.  
J. Bacteriol. 164, 639-645, 1985  
A|Title: Structure of two divergent promoters located in front of the gene encoding pull  
A|Reference number: A91807; MUID:86033621  
A|Accession: A25025  
A|Molecule type: DNA  
A|Residues: 1-65 <CHA>  
A|Cross-references: GB:M12503; NID:g149223; PIDN:AAA25087.1; PID:g551870  
R|Kornacker, M.G.; Boyd, A.; Pugsley, A.P.; Plaszow, G.S.  
Mol. Microbiol. 3, 497-503, 1989  
A|Title: *Klebsiella pneumoniae* strain K21: evidence for the rapid secretion of an unacy  
A|Reference number: S04864; MUID:89345636  
A|Accession: S04864  
A|Molecule type: protein  
A|Residues: 'X', 21-31, 'X', 33, 'X' <KOR>  
A|Experimental source: strain K21  
C|Genetics:  
A|Gene: *pula*  
C|Keywords: glycosidase; hydrolase; lipoprotein; polysaccharide degradation  
F|1-15|Domain: signal sequence #status predicted <SIG>  
F|20-1038|Product: pullulanase #status experimental <MAT>

Query Match	78.7%	Score	4435.5	DB 2	Length	1038			
Best Local Similarity	83.3%	Pred.	0.5.6e-261						
Matches	873	Conservative	52	Mismatches	84	Indels	39	Gaps	10

```

0Y      1 MRYRYRNALVIGSLVLLSGCNGSSSSSSGQNP- ---TPQNDQVYVRLPDVAVNGEAVTA 57
Db      1 MRYRYRNALVIGSLVLLSGCND-SSSSSSGSPQNPQNDQVYVRLPDVAVNGEAVMA 59
0Y      58 VENOAVIHLVDIAGITSSADYSSKKNLYLNNNETCALSAVADAMNDVSTTPGSGDRYG 117

```

Db	60	TANQAVIHLVDINGITGSTPDSXNKILNMNENCODALSAPVADMNVDSTTPAGSDKYG	119
QY	118	PIWVPIPLNKSSGCIINVIYRDGTDKILDSLRVAFGDFDRIYSVIAGNSAVYDSRADAFR	177
Db	120	PYWVPIPLNKSSGCIINVIYRGSTKILDSLSLRVSFGDFDRIYSVIAGNSAIDSRADAFA	179
QY	178	AAPGVALAEHWADKNTLILMPGODKPIYRLVYSHSSKVAAGDEKFTDRIYKILPTTVS	237
Db	180	PLSAGRLPKLTGLIKRRCC-GRGEDPIYRLVYSHSSKVAAGSEKFTDRIYKILPTTVS	238
QY	238	QOVSMPREPLHSYAAFKLPDMMANVDELLQGETVAIAAAEDGILISATOVOTAGVLDDAYA	297
Db	239	QOVSMPREPLHSYAAFKLPDPAKTDELLQGETVAIAAAEDGILISATOVOTAGVLDDYTA	298
QY	298	EAAALSYGQALDGGVTFPFWAPVPAOQVDVYVYSADKKVYIGSHWTRDSAGSMYOGG	357
Db	299	PAALALVYGQVSDGGVNFELMAPVPAOQVDVLYYSADKKVLAISHWTRDSTSGSWTMOGG	358
QY	358	SDLGAFYRYAAMVYVYHPOSKRYEQEYVETDTPYAHLSISTNSEYSOQVVDLNDLSALPPGMNL	417
Db	359	SDLGAFYRYALYVYHPOSKRYEQEYVETDTPYSHLSISTNSEYSOQVVDLNDLSALPPGMNSL	418
QY	418	TMFHAQTKTADLAKMTIHESHIRDSLSDMDQVPAELRGKYALTLAGDSNMVYOHILKTLISAS	477
Db	419	TMFHAQTKTADLAKMTIHESHIRDSLSDMDQVPAELRGKYALTLAGDSNMVYOHILKTLISAS	478
QY	478	GVTHEVLLPVPFDLATVNEFSDKVADIOOPFSLCEVNSVKSSEFAGYCDGSGTYAEVLN	537
Db	479	GVTHEVLLPVPFDLATVNEFSDKVADIOQSSHLCEVNSVKSSEFAGYCDGSGTYAEVLN	538
QY	538	QLKQSDSDQNDPOVALNTLTAQDTSYNMGIDPEHYVYVPGSVAATDPEGTRIKERFTMIQ	597
Db	539	QLKQSDSDQNDPOVALNTLTAQDTSYNMGIDPEHYVYVPGSVAATDPEGTRIKERFTMIQ	598
QY	598	AIKQDCLMNTIMDVYVYHNTTAAAPRTORTSYLXKIVYVYVQRLNETTGSYESATCCSDSAP	657
Db	599	AIKQDCLMNTIMDVYVYHNTTAAAPRTORTSYLXKIVYVYVQRLNETTGSYESATCCSDSAP	658
QY	658	EHRHFAKLIADSLAVWTTDVKIDGFPFDLMGYHPKQOILISAMERIKALNPDIYFFGEGMD	717
Db	659	EHRHFAKLIADSLAVWTTDVKIDGFPFDLMGYHPKQOILISAMERIKALNPDIYFFGEGMD	718
QY	718	SNQSDRELIASQINLAKGTGIGTFSDRLRDSYKGGGFPDSDGLR--QNOGIGSGAGVLPN	775
Db	719	SNQSDRELIASQINLAKGTGIGTFSDRLRDREPCAAA--RSIPAMRCVTKRGLAAA-----R	771
QY	776	ELASIDDOYRHLADL-----TRLGAMGNLADFMYIMIDKGAAKGSEIYVNGAPGYADP	831
Db	772	RVAKRADHPERRPSSSPGRPDSSGDRNLADFLIIDKDAVKKGSEIYVNGAPGYADP	831
QY	832	TEVVYVYSKHDNQTLMMDISYKASOEADLATRVYMOAVSLATVYMLGOGIAPFOQGSSELLR	891
Db	832	TEVVYVYSKHDNQTLMMDITYKAAQOQADLWTRBYMOAVSLATVYMLGOGIAPFOQGSSELLR	891
QY	892	SKSTPRTSYSGDMFENVDYSIODNNYNYCMGPRISDGSNTEYIVTRVKEKMATPREAEKL	951
Db	892	SKSTPRTSYSGDMFENVDYAMQDNNFNMGMPISDGSNTDYIVISRVKEKMATPREAEKL	951
QY	952	QMTAFYQELTELKRSQPLFTLJGDSAVMKRVDRNTGSPQOAGILLVMTVD-----1001	
Db	952	QMTAFYREIETGLKRSQPLFTLJGDSAVMKRVDRNTGSPQOAGILLIMTIDMACRQGRVSI	1011
QY	1002	QMKAGASLDSRLDGLVVAIINAAPESRT	1029
Db	1012	AGLMTGGS-DOR-----ASESRT	1028

RESULT 4  
S51324  
pullulanase - spinach  
C:Species: Spinacia oleracea (spinach)  
C:Date: 07-May-1995 #sequence\_rev:101 01-Sep-1995 #text\_change 09-Sep-1997

C:Accession: S51324  
 R:Renz, A.; Schmid, R.; Kosmann, J.; Beck, E.  
 submitted to the EMBL Data Library, January 1995  
 A:Description: Sequence of a cDNA of the Splancla oleracea L. starch debranching enzyme.  
 A:Reference number: S51324  
 A:Accession: S51324  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-964 <REN>  
 A:Cross-references: EMBL:X83969; NID:g634092; PID:g634093

Query Match 26.7%; Score 1503; DB 2; Length 964;  
 Best Local Similarity 36.2%; Pred. No. 4.2e-83;  
 Matches 365; Conservative 169; Mismatches 324; Indels 150; Gaps 33;

123 PLKESGGINVIRDGTDLKIDSLVAFGDFDRIVSYAGNSAVYDSRADAFRAFGV 182  
 53 PLKDDSC-----PCCSMAYEVGSASSVSOSELOGS-----L 84  
 183 ALAEAHMVDKNTLLW---PGGDKPIVRLYSSSKVADGEGKFTDR-----YLKL 231  
 85 NSCRATWPSKYTFAMVNDIGNS-----YLFASKTA--LKFTDAGEGYDKIKL 134  
 232 -----TPPTVSQOVSMRPHLSYAFKLPDANVDELQGETVAIAAEDGILISATQ 285  
 135 DKDGGILPAAVTE---KPHINGYSAFKAPATLDVDSLKCOLAAVAFSADACRNATG 190  
 286 VOTAGVLDADYAAEAALSY---GAQLADGVTFRVMAPTAOOVVVYSAKKYISG 341  
 191 LQIPGVIDEY-----SIDGPLCAVFSNNTSLYLMATPAOVASISIF---KDSGGE 240  
 342 PMTR---DSASGAMWOGSDLKGAIFYRYAMTVYHPOSRRVEGYEVTYPAHSLSSTNEX 398  
 241 PLQTVOLIESNGVSAVAGPRTMEGCVYVEITYYHSTLIEKSFADIPYARISADYKR 300  
 399 SOVVDLNDSLKRDGNDNL--WRHAKTKADLAKMTIHESHRLDSAMQUTYPAELRCK 456  
 301 TLLADLSSETLKEGEMENLADKPHLSP---SDISLELMIRDSAVADLTYPDLRG 356  
 457 YLALTAGDSNMVOHLKTLASGVTHVELLPVFDLATVNEFSDDKVDIQOPFSRLCEVNSA 516  
 357 YLAFITDSDSGVHLEKLSAAGLTNHLILPSPFAVEDD--DK----- 397  
 517 VKSEFAGVCDGSIVEEVLNQLKSDSDNPVOVALNTLVAOTDSTYMGVDPFHYTPE 576  
 398 -KKMKF-----VDTKREFELPRDSEGOAQTIRDEGDGYNMGVYNPVLMGTPK 444  
 577 GSYATDPEGTRIKERPTMIOALKODLGMNVINDVYVNHNAAGPDRGTVLDKIVPWWY 636  
 445 GSYATDNGPCRIIEFKMYOALNR--IGLRVYLDVYVNHLSGSPSDNSVLDKIVPGY 503  
 637 QRLNETTGSVESATCCSDSAPRHRMFAKLADSLAVTYYKIDGFRFDLGMYPKAOIL 696  
 504 LR-RDNGCALENGCVNDTASEHFWERLLTLDLKHMAVYVYKDGFRFDLGMIMKTIV 562  
 697 SAMERIKALNPD-----TYFFEGWDSNQ---SDREFIASQINLKGTGIGTFSDRLR 745  
 563 KATNMLOGLSKNIDYEGSSITLYEGEWMDEGEVANNARAGVNASQINLGCTGIGSFNRR 622  
 746 DSVRCGPFGSCDALRNOIGSAGVLPNE---LASLSDOYRHILA-DLTRIGMGNL 801  
 623 DAVLGGPR--GPRLO--QGYVTGLSLQPNDDHDSKANADRMALIAKKDIIQVGMAGNLR 678  
 802 DFVVIDDGAAGKSGSEI-DYNGAPGYAADPTEVYVNSKHQNTLMDMTSYKASQEADL 860  
 679 DYLLTNDGQOVGSEYVYTGCPYVYAMQPIETINVSAHNDETLFDIYSLKTPYITV 738  
 861 ATVRMQAVLATVMLGGIAPDOQGSLLRSKSFTRDSYDSGDWFRVVDYSLQDNNYV 920  
 739 DECRVYHNLTLSTLALSGQIPFTIAGDELLRSKSLDSDSYNSGDMWRDLDFSYSNMGV 798  
 921 GMRISIDGDSNRYIVRYKEMVATPG-EALIKQMTAFYQELT---ELRKSSPLFTLDGS 976

DB 799 GLPPKDHNSMPLI---KKRLANPSYKDPKNIHAIVENFTMLQIRYSSPLFRLSAK 855  
 977 AVWKRYDFRRTGSDQAGLLVMTVDGKAKA-----SIDSRLDGLVAVYINAAAP-ESRTLN 1031  
 DB 856 DIEDRVRFHNNVPSWIPGLANSTIEDG-HGAGPGLSOLIDRKFQYIYVILNVOPTETKPVN 914  
 1032 -EPAGETLQLSAIQOTAGENSLANGVOIADGTVTLPAWSVAVLELPO 1078  
 DB 915 PDLRAKSLQJHPVOSTSGDTPVVEKSEKPESTGCTIPRSTAVFEPR 962

RESULT 5  
 101688 starch debranching enzyme, pullulanase-type - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 29-Oct-1999  
 C:Accession: T01688  
 R:Beatty, M.K.; Rahman, A.; Cao, H.; Woodman, W.; Lee, M.; Myers, A.M.; James, M.G.  
 Plant Physiol. 119, 255-266, 1999  
 A:Title: Purification and molecular genetic characterization of ZPUL, a pullulanase-t  
 A:Reference number: 214395; MUID:99097061  
 A:Accession: T01688  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-962 <BEA>  
 A:Cross-references: EMBL:AF080567; NID:93411265; PID:AA011599.1; PID:93411266  
 A:Experimental source: endosperm  
 C:Genetics:  
 A:Gene: ZPUL  
 C:Function:  
 A:Description: hydrolyzes alpha 1,6 glucosyl linkages

Query Match 25.4%; Score 1432.5; DB 2; Length 962;  
 Best Local Similarity 35.7%; Pred. No. 8e-79;  
 Matches 336; Conservative 173; Mismatches 338; Indels 93; Gaps 24;

181 GVALAE-----AHWVKNTLLPVGQDKPIYRLYSSSKVAAADGEG-KFTDRYL 229  
 73 GVAVESAOGFLDARAYVWTKSLIAMNISDKTSLFLYASRNATMCMSSQDKGYDSKY 132  
 230 KLPPPT--VSQOVMRPHLSYAFKLPDANVDELQGETVAIAAEGCILLISATQV 287  
 133 ELDPENDGLPSSVTKQKPFSSYRAFERIPSSVDATLVKCOLAVSDAHGNRODYGLQ 192  
 288 TACVLDAYAEALSYGQALADGVTFRVMAPTAOOVVVYSAKKYISGHPTR-- 345  
 193 LPGVLDMMFAYTG---PLGITFSEAVSNYLMAPTADVSFYTDP-----AGPLLETV 244  
 346 --DSASGAMWOGSDLKGAIFYRYAMTVYHPOSRRVEGYEVTYPAHSLSSTNEXSYVD 403  
 245 QLMELNVMGVSVPTRMNNENYUYLEVTVYHQTGNEKCLAAAPYARGLSANSTRMLVD 304  
 404 LNSALKRPGDMNLTJPHAKTKAD-LAKMTIHESHRLDSAMQUTYPAELRCKYALTA 462  
 DB 305 INNETLRLAMDGLA---AKRPRDSFSDISTELHRLDSAHSDSTVDCPRRGFCATF 361  
 463 GDSNMVOHLKTLASGVTHVELLPVFDLATVNEFSDDKVDIQOPFSRLCEVNSAKSSEF 522  
 362 QDSVGLIELKLKSLDAGLTNHLILPSPFGGVDD-----IKS--- 397  
 DB 523 AGYCDGSSTVEV-LNOLKO-SDSQNPVOVALNTLVAOTDSTYMGVDPFHYVPEGSYA 580  
 398 -----NMKCYDELELSKLPGSDIQ-----QALVALIOEDPYNMGVNPVVMGVPKGSYA 447  
 581 TDPBGTTRIKERPTMIOALKODLGMNVINDVYVNHNAAGPDRGTVLDKIVPWWY 640  
 448 SNPGPSRIETIRLMYOALNR-LGLRVVMOVVYVNHLSGSPATTSVLDKIVPGYTLR-R 505  
 641 ETTGSVESATCCSDSAPRHRMFAKLADSLAVTYYKIDGFRFDLGMYPKAOILSAME 700  
 DB 506 DSNQGTENSAVAVNNTASEHFWERLLIYDILLMAVNVYKVGDFRFDLGMIMKTIVAKS 565

QY 701 RIKALNPP-----IYFEGMDSN---OSDRELIAOS:INLGTGIGTFSPLRLSDSVR 749  
 Db 566 ALQSLTIDEHGVQDSKITYLGEQWNFGEVAENQKINGCSQJLMSSTGIGSFDRIRDAIN 625  
 QY 750 GGGPFDSGDALRQNGIGSGAGVLPNELASLDDOVR---HLADLTRLGAGNLADEVM 805  
 Db 626 GGSFP--GNPLQ--QGFSGTGLEPNNGFYQGNETFTRLTLATVYDHIQIGLAGNLKDYV 681  
 QY 806 IDKGAAKKSGEI-DYNAPGGYAADPTEVYVNYYSKHNOIIMDMISKASQEADLATRY 864  
 Db 682 ISHGEARKGSEIRFTFDSPPGVYASSPIETINYASAHNDNETLFDLISLKTPTDLSIDRC 741  
 QY 865 RMOAFLATVMLGOGIAPDOOGSELLRASKSFTROSDYSDGDMFNRYDYSLODNNYVGM-P 923  
 Db 742 RINHLSTSMIALSGIIPFFHAGDELRLRSKSLDRQSDYSDGMFNKIDFTYETRNMGVGLRP 801  
 QY 924 RISDDGSNEYITRYKKEVATPGAEELKOMTAFYOEITELRKSSPLFLTGDSAVMKRVD 983  
 Db 802 REKNKGSWPLMKPRELNSFPKAKHDIILALDKFIDILKIRYSSPLFLRLTASDIAVGRVH 861  
 QY 984 FRNNGSDQAGILVNTVDDMKAG---ASLDSRLDGLVVAIINAA-P--ESRRLNEFAGETL 1038  
 Db 862 FHNNGPSLVPGYVWISIDARNDHRDMAQIDETFSCVTVFNVQCYEVSIEIPLDLSLRL 921  
 QY 1039 QLSAIQQTAGENSILANGVOIADGTVTLLPAMSVALLPQ 1078  
 Db 922 QLNHPVQVNSDALARQSAVYDTATGRFTVPRKRTAAVAFVEPR 961

RESULT 6  
 T04377  
 Probable alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) - barley  
 N:Alternate names: pullulanase  
 C:Species: Hordeum vulgare (barley)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_Change 08-Oct-1999  
 C:Accession: T04377  
 C:LOK, F.: Kristensen, M.; Plancho, V.; Leah, R.; Svendsen, I.; Svendsen, B.  
 Submitted to the EMBL Data Library, December 1997  
 A:Description: Isolation and characterization of starch debranching enzyme, limit dextrinase  
 A:Reference number: Z15320  
 A:Accession: T04377  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-904 <LOK>  
 A:Cross-references: EMBL:AF022725; NID:g2502057; PIDN:AAD04189.1; PID:g2677837  
 A:Experimental source: cv. Igrl  
 C:Genetics:  
 A:Gene: HvIDP9  
 A:Introns: 21/3; 70/1; 87/3; 124/1; 169/3; 235/2; 285/2; 305/3; 346/1; 376/2; 407/3; 430/1  
 C:Keywords: glycosidase; hydrolase

	Query Match	25.3%	Score 1424.5;	DB: 2;	Length 904;	
	Best Local Similarity	35.7%	Pred. No. 2.2e-78;			
	Matches 330;	Conservative 176;	Mismatches 341;	Indels 77;	Gaps 23;	
Oy	185	AEAHVNDKNTLLMPGQ--DKPIVRLYYSHSKVA--ADGECKFMDRYLKLPTT--VSQ	238			
	:     :	:   :   :	:   :   :	:   :   :	:	:
Dd	27	ARAYVVTSDLIIMNNGELEAQSVCLYASAAAMSLSPSGIGGYDSKVELOPESAGLE	86			
Oy	239	OVSMPHILSSYAFLKLPDNNANDELQETVAIAAEDGILLISATQVOTAGVLDAVAE	298			
	: :     :           :	: : : : : :	: :   :         :			
Dd	87	TYTQKFPTISSRAKRVSSVDVALYVCQLVVASFGADKHVDVTLGLPVLIDMFAY	146			
Oy	299	AAELNSYQAQLDGGVTFRRVAPTAQQDVVVVYSADKKYIGSHPMTRDASGAWMSGGS	358			
	:   :   :   :           :	:   :   :   :   :	:   :   :			
Dd	147	TG---PLGVVSEDSVSLHLMLAPTLAGVSVCFDDPGAPALFTVOLKE--NWWSVTGR	202			
Oy	359	DLKGAFYRAMTVVYHPQRKVRQOYEVTTPYAHSLSTNSEKYSOVVDLNDALKPPDGWDNLT	418			
	: : :   :         : :	:   :         : :	:   :   :   :       :			
Dd	203	EEMENNYLYLEVVDVYIPTKAQVYLKCLAGDPYASLSANGARTLVLDINNTELTAKPASWDELA	262			

```

0Y 419 MPHOKTAD--IAKTHIESHRLDAMDOVPAPELREKCTALATPAAGSNMVOHLETKISAS 477
Db 263 ---DEKPLDLSFSDDITTYIELHIRDFSAHDGTYVDSKRGCFAPAYOASAGHEHRLKLSDA 319
0Y 478 GVTHVELLPVFDLATVNEFSDKVADIDQOPFSRLCEVNSAVKSS--EFACYCDSGSTVEEVL 536
Db 320 GLTHVHLLPSEHFAGVDD-----IKSNKFEVDEC-----ELA 351
0Y 537 NOLKOSDQDNROVALMTLVAQOTDSYMWGYDPETHYVPEBSYANDPCTTRIKEFRPMI 596
Db 332 TFPQSDMQ-----OAAVVAIQEEDPTNMGNPVALMGPKOSTYASDPGSPRILLEYROMV 406
0Y 597 OAIKODLGMNVMIDVYVNHNTNAAGPTDRTSVLDKIPVYQORLETTGTSVSATCCSDSA 656
Db 407 QALNR-ICLRVYMDVYVNHLDSDSGPCGISSVLDKIVPYVR-RDINOQIENSAAMNNTA 464
0Y 657 PEHRFALLIADSLAVMTTDKIKGFRFDLMGHRPKAOILSAMERIKALAND-----708
Db 465 SEHEFVNDRLIYDDLLNMAVNYKVDGFRFDLMGHIMKRTMVTK-SALOSLTJTDAGGVDSK 523
0Y 709 IYFEGEGDMSQ---SDRFELASQINLKGTCIGTFSDRLRDSVRGCGPFGSDGALRONOG 765
Db 524 IYLVGEGNDPHEVANRONGINGINSQILMSGTCIGSPNDIRALINGNF--GNPLQ--QG 579
0Y 766 IGSAGVLPNELASTSDQVRH----IADLTRLGAGNLADPVMIDKDGAAKKGSEI-DY 820
Db 580 FNTGTFLEBPNGFYOGNEADTRSLATYADQIOIGIAGNLARDYVLLSHGGEAKKSEITTF 639
0Y 821 NCAPGCVAADPTEEVNYSKHDNOFLMDMISYKASQADLMTNRYMVAQVSLATYWLGGOI 880
Db 640 DGLPAGYASPIETINYSAHNDNETLFDVYSKTMILTSVERCIRINILASMAALSOGI 699
0Y 881 AFDOOGSEILNSKSFTRDSYDSGDMFENRVDYSLQDNNTNVGMPRISDDGSNEYEI-TRVK 939
Db 700 PFFHAGDEILSKSIDRBSYNSGDWFNKLDPETYETNNMGVCLPSEKNEEDNMPLMKPRLE 759
0Y 940 EVVANPGEAEIKOMTAFYQOELTELKSSPLRTLDGSAVYMRVPERNRGSDQOAGLLVMT 999
Db 760 NPSFPAKGHILALDSTVDYDIKITYSSPLRLSTANDIKORVHRHNTGPSLVRGVYMG 819
0Y 1000 VDDGKAG--ASLDSRLDGLVAINAPE--SRTLNEFAGETILOLSAQTAGENSIAN 1054
Db 820 IEDAGESPENAOJLDTFNFSYVTVENVCPEHYESMDIPALASMGFELHVPYVNSSDTLVRK 879
0Y 1055 GVQIADAGTVTLPRMSVAVLELPQ 1078
Db 880 SAYEAAATGRFTVPGRTVSFVEPR 903

```

```

RESULT 7
75524
Alpha-dextran endo-1,6-alpha-glucosidase - Delnoccocus radiodurans (strain R1)
:Species: Delnoccocus radiodurans
:date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
:Accession: D75524
:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.R.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
: Smith, H.O.; Venter, J.C.; Fraser, C.M.
:Science 286, 1571-1577, 1999
:Title: Genome sequence of the radioresistant bacterium Delnoccocus radiodurans R1.
:Reference number: A75250; MUID:20036896
:Status: preliminary
:Accession: D75524
:Molecule type: DNA
:Residues: 1-910 [SMU]
:Cross-references: GB:AE001900; GB:AE000513; NID:96456079; PIDN:AAF09983.1; PID:g645
:Experimental source: strain R1
:Genetics:
:Gene: DR0405
:Map position: 1

Query Match      25.0%   Score 1405.5;   DB 2;   Length 910;

```

Best Local Similarity 35.3%; Pred. No. 3.2e-77;  
Matches 345; Conservative 157; Mismatches 365; Indels 109; Gaps 27;

131 INVYDGTDLIDSLRVAFGDET---DRTVSIAGNSAVYSRADAFRAAG-----181  
8 LGFTVHKGDGDKDPD-----GTLTRGQGNQIMIVSGNPSTTRDTRVRYGDLKQOQ 62  
182 -VALAEAHVYDKNTLLPGCGQKPIRYLYSHSS--KVAADGCKFTDRYLKLTPTTVSQ 238  
63 AILSLSDLLYVVKPALAQPGAM---LTLHVSPTATLKLGDVSGSDTLALEVEGGLTP 118  
239 QVSMRPHLSVYAPFL--PDNANVDELLQGETVVAIAAEDGILLISATQVQTAGVLDAY 296  
119 ALQAKAPYLAGYLLQIRADRAKKLPTALSGLAVSSVLPQGVYLDVQYVAMALDDL 178  
297 AEAALSTGAQIADGCVTFRRVAPTAQOVVVYVYADKKVIGSHMPTRDSAGSAGMSQG 356  
179 AVAG---PLCVTWQGNKPIRYLMAPTAQOVYKLVSRPDQGETRLMTROT--QGVMTASG 234  
357 GSDLKAFRYATVYVHPOSRAKEYEYVDPYAHSLSTSEYQVYVLDLSDALKPDGMWN 416  
235 DQSMKGLAFREFEYVAPSTGNETNLVYDPSVALTRNTHSVFADLSNSOKPAGMNA 294  
417 LTPHQAOKTKADLAKMTIHESHIRDSAMDQYVPAELRGKYLTLFAGDSNMVQHLKTL 476  
295 LKRP-ALRSPADL---SPYELHLDPSADASVPAABRGTYLFTPRGSDGMHLALAG 350  
477 SCVTHVELLPFDLATVNEFSKVDADIOQFSLCEVNSAVKSEFAGYCGSGTVEEVL 536  
351 AGLKAVHLPTFDIATINE-----DKQW--QTP 377  
537 NOLKQSDSDNPOVALNTLVACTDSYNNQYDPFHYTVPEGSATPDEGTRIKERFTMT 596  
378 GDLTQCGPNSEEQKAVNA--VKDADAVYNNQYDPYHSMVEGSAVNP--AERKERYOMV 434  
597 QAITKODLGNVINDVYVYNNHTNAAGPTDRTSVLDKIVPMWYQRLNENTGSEVATCCSD 656  
435 MSJMA-AGLRVQDVYFNHTMAAGLAE--SVLDKIVPGYTHRLN--VNGCYENSTCCSMT 491  
657 PEHMFPAKLADSLAVMTDYKIDGFRFDMGHPRAQILSAMERIKALNP-----D 708  
492 TEHTMRRLMVDTLVLMAREYKVDGFRFDMGHMVAQMAARALDALTPAKDVGDKQ 551  
709 IYFPGCMQSNQ---SDREIASQINLKGICGTFSDRLRDSVRGCGPDSGALAKQNG 765  
552 IYLYGEGMDEGEVQGNRKGANATQNLVYAGIGTFENDRVADALRGSPFGG---LQEG 607  
766 IGSAGVLPN-ELASLSDQVRHLADLTRLGMAGNLADFMVIDKDAAKGSEIDNGAP 824  
608 VAGGLATVPNGCAGNSAEKMGQLADLVKLGAGNLRDYRFTDHSQKPYVGGQLRGDAP 667  
825 GGYADPTEVYVYVSKINDOTLMDMISYKASQEADLATRYRMOAVSLATVYLQGIAPDQ 884  
668 AGTAASPRETINASADHNOTLMDAVLTKTPAMNSAARVRQNLASHYLLIQGLPFPV 727  
885 QGSELLRSKSFPTSDYSDGDMENRVDSYLDNNYVNGMPRIISDDGSN---YEVITVKE 941  
728 AGGELLRSKSFPTSDYSDGDMENALDMDTGTITNGFGKGLPRAEKENEMMALYRTLLD 787  
942 VATPGEAELOKMTAFYOELELRLKSSPLFTLGCGSAYVMKRVDRNNGSDQOACLVTVD 1001  
788 KYVP--ADITRASDHRELLRVRSSSLFRLDTAQVOOSLSF-----LAPTYIAMK-- 839  
1002 DCKKACASLDSRLDGLVVAIINAPESRTINEFAGETLQISAIGOTAGE-----NSL 1057  
840 --LSGVSATNPYRDLVITFNGSADAVTLINDASAGMKLHPVLAASSDPVIRTSKVSQ 897  
1058 IADDTVTLPAMSVAV 1073  
898 -----ASVPALTAV 907

RESULT 8

T03760  
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) - rice  
N:alternate names: pullulanase  
C:species: Oryza sativa (rice)  
C:date: 23-Apr-1999 #sequence, revision 23-Apr-1999 #text, change 21-Jul-2000  
C:accession: T03760  
R: Makamura, Y.; Umemoto, T.; Ogata, N.; Kuboki, Y.; Yano, M.; Sasaki, T.  
Planta 199, 209-218, 1996  
A:title: Starch debranching enzyme (R-enzyme or pullulanase) from developing rice end  
A:reference number: Z15063; MUID:96275956  
A:accession: T03760  
A:status: preliminary; translated from GB/EMBL/DBJ  
A:molecule type: mRNA  
A:residues: 1-986 <NAK>  
A:cross-references: EMBL:D50602; NID:g836661; PIDN:BA09167.1; PID:g1783306  
A:experimental source: cv. Fuhikari, endosperm  
C:genetics:  
A:map position: 4  
C:keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 23.5%; Score 1326.5; DB 2; Length 986;

Best Local Similarity 33.8%; Pred. No. 2.2e-72;

Matches 333; Conservative 160; Mismatches 360; Indels 131; Gaps 26;

164 GNSAVYDSRADAFRAAFGVALAEAHVYDKNTLLPGCGQKPIRYLYSHSSKY-AADGEG 222  
64 GRDVGEECAAAVAGSGFVTDARAVYTRSLIMANNVDQTSFLVASRDATMHSVDGAI 123  
223 KFTDRYLKLT--TYSQOVSMPHLSYAPFLPDNANVDELLQGETVVAIAAEDGIL 280  
124 HGYSKIELEPEHASLPDVAEKEFPIRSYRTPRVSVDVAVSLQALVASYADAHGRN 183  
281 ISATQVQTVGVLDDAAEAALSYGAQLADGCVTFRRVAPTAQOVVVYVYADKKVIGS 340  
184 QDVTGLQLPQVDDMRAVTC---PLGAVPSDKVDYLMAPTAQOVRYCVDP-----A 225  
341 HPMTN-----DSAGSAGMSQGSGLKAFYRYAMTVYVHPOSRAKEYEYVDPYAHSLSTNS 396  
236 GPLQTVQLKELGNVSVTVPRYRENOYLYLEVKYVHNPSTSYQVEKCLADDPYARGLSANG 295  
397 EYQVQVVDLSDALKAPGQWDLTPHQAOKTKADLAKMTIHESHIRDSAMDQYVPAELRGK 456  
296 TRTWLVDINSETLKPASWDEL--DEEPLSEPSDSTIELHLDPSADHSTVDKNSRG 353  
457 YIAL-----TACDSNMVQHLKTSASGVTHVELLPYV 488  
354 FVHLHFRFLRLNLNDFCSPPIRKHPGRIMETVMQDSAGIRHLRKISAAGLTHVHLPSF 413  
489 DLATV--NEPSDKVADIOQFSLCEVNSAVKSEFAGYCGSGTVEEVLNOLKO--SDSQ 545  
414 HFASVDDNKSMMKFDV-----EADLAKLPDSDEQ 443  
546 DNPQVALNTLVQPTSYNNQYDPFHYTVPEGSATPDEGTRIKERFTMTQAIKODLGM 605  
444 ----QAATYSIOEDPYNMGYDPLVMGVPKGYASAPDPSRIETRQVQVLANR--TGL 497  
606 NVIMDVYVNNHTNAAGPTDRTSVLDKIVPMWYQRLNENTGSEVATCCSDAPERRHFAKL 665  
498 RYVADVYVNNHLDSSGPGVSSVLDKIVPGYLLRN--VNGIENSAMNNHTASHFVNDRL 556  
666 IADSLAVMTDYKIDGFRFDMGHPRAQILSAMERIKALNP-----IYFPEGGMD 717  
557 TVDDLLMNAINTYKVDGFRFDMGHMKSMTIRAKSAIRSLTRDVHGVYSKIYLYGEGMD 616  
718 SN---OSDREIASQINLKGICGTFSDRLRDSVRGCGPDSGALAKQNGIGSGAGVLP 774  
617 PGEVAGNKKGINASQILMNGSTGISFPNDKIRDSVNGNRP--GNPLQ--QGFSTGLFLEP 672  
775 NELASLSDQVRH---LADLTRLGMAGNLADFMVIDKDAAKGSEI--DYNAGAPGYAA 829  
673 NGYVQGNREADTRRELATYADHIQIGLAGNLKDVLTHTHGEAKKGSDIYTFDSSPVGYS 732

[illegible]

```

RESULT      9
H72204      \
pulinulanase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: H72204
R:Neelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: H72204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-843 <ARN>
A:Cross-references: CB:AE001821; CB:AE000512; NID:94982421; PIDN:AMD36907.1; PID:94982424
A:Experimental source: strain MSB8
C:Genetics:
:Gene: TMI845

```

Query Match	12.5%	Score 703.5	DA 2	Length 843
Best Local Similarity	24.9%	Pred No. 1.1e-34		
Matches 259	Conservative 150	Mismatches 386	Indels 247	Gaps 39

[illegible]

Dh	323	NPBGEWMDRQPKIEGVEDAI----	YIEIHADITGLENS-GVKNNGVLTGLTEWTKGCG	377
Qy	466	MMVOHLKTLASGYTHAVELLPV	VDLATVNEFSSKXVADIOQPSRLCEVNSAVKSSFEAGY	525
Dh	378	GVTTGLSHLVELGATVHHHLL	PFEDFYTGDEL-DK-----	DFEY 415
Qy	526	CDSGSTVEEVLNQIKQSDSDQ	NPQVALMTLVAGTDSYNNMGDPFHYTVPEGSATQDEG	585
Dh	416	-----	YMKGDPTLFMVPBRYSTDKN	438
Qy	586	T-TRKEFRMTQAIKODLGMN	IMDVYNNHTNAAGPTORTSVLQIKVPMYTORLNETTG	644
Dh	439	PHTRIRREVKEMVVALNHK	-GIGVIMMVEPHTGIG---ELSAFDQTVAYEYRLDKTGA	494
Qy	645	SVEASATCCSDAPEHRFA	LIALDSLAVMTDYKIDGPFREDLGMGHPKQOILSAMERKA	704
Dh	495	YLNSSGCGGNVYASERP	PMKRFYIDVTYVYKEHHIDGFFRFDQGLIDKTKTILEVERALHK	554
Qy	705	LMPDIYFPGGWDMS-NQSD	RFEIASQILNKGTGICTPSDRLRDSVNGGSPFDSGALRON	763
Dh	555	IDPTIILYGPBGWGPAPIRF	---GRSDVAGTHVAAFDEFBDAIG-----	598
Qy	764	QGTSGAGVLPNELASISDQ	VOVHLADLTRLGMAGLADPYM-IDXDAAKG--SEID	819
Dh	599	-----	SVFNPSVK-----GFWMGYKKEIRKRGVGSIN	628
Qy	820	YNG-APGAYADPTEVYNN	YSKHDNQTLM-D-MISYKASQADLADTRVMQAVSLATVWL	876
Dh	629	YDGLKLSFALDPETINT	YNAACHDNHTLMQXNLAKKAKKREWEETELKNQOKLAGALL	688
Qy	877	---GGGIAFDQOQSGSL	SKSFTRDSDSCDMPKRVDTSLQDNNTYVNGMGRISDQGSNTEV	934
Dh	689	LTSQGFVFLHGGQDFCTK	TK-----NDNSYNA--PISINGDYDE	726
Qy	935	ITRKEMVATPGEAELE	KOMAFYOEELTERKSSPLTELDGSSVAMKRVDFRTGSDQOAG	994
Dh	727	-----	RKIQFIDVYNHNGKLKLKKEHPARLKNABEIKKHLEFLCGGRIVAF	775
Qy	995	LLVMTVDDGKAKAGAS	LDSDRLDGLVVALINAAPEBRTLINEFAGETLQLSAIOQTAGENSLAN	1055
Dh	776	ML-----	KDHAGGSPKMD-IVVIYNGMLK-----TTYKLPBECKMMVNVNSOKA	818
Qy	1055	GVOL- AADGTVTL	PAMSVAVL 1074	
Dh	819	GTEVIEIVEGTIELOP	LSAIVYL 840	

RESULT 10  
G69585  
pullulanase amyX - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence-revision 05-Dec-1997 #text-change 15-Oct-1999  
C:Accession: G69585  
R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bevilacqua, L.; Broer, S.; Bröllert, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C. Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall, Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulo, M.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau, S.; Ogawa, K.; Ogilwara, A.; Oudgaa, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scorfione, F.; Sekiguchi, J.; Sekowska, A.; Sekouchi, M.; Tamakoshi, A.; Tanaka, T.; Terpilata, P.; Tognoni, A.; Tosato, V.; Uchida, T.; Winters, P.; Wipolt, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
A:Reference number: A69580; MUID:98044033  
A:Accession: G69585  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
Molecule type: DNA  
Residues: 1-718 <KUN>



QY 788 LADLRLGMAGNLADFWIDKGAAGKSEI--DYNAGPGY--AADPTEVYVYKSHDNQ 844  
 Db 754 IT-----GGARNIEQJFDNLTAQPHNETATNPEDVPIYEAHNL 793  
 QY 845 TLMNMISYKASO-----EADLATRVMOAVSLATVMLGCGIAPDOGSELLRSKS----- 894  
 Db 794 TLHDVIAOSIRKREYHOEEIHKRIRL--GNTVVLTAQCTAFIHAAGEFRTQOFADT 850  
 QY 895 -----FTPDSDGDMENRVDSLODNVAVGMPRIISDGSN 931  
 Db 851 GGAAPYKSTYMTDENGKRFVYPIFHDYSDDSAINKFKDMKATNK-----EK 898  
 QY 932 YEV--ITRVKEMVATPGEAEIKOMTAFYQELTELKSSPLETLG--DGSAYMKRV-- 983  
 Db 899 YPIHHLR-----TYTAGIELLRSTNATFTHGLTVQVQANVRLDDE 941  
 QY 984 -----FRNTGSDQAAGLLVMTVDDGMAKASL--DSRLDGLVY 1019  
 Db 942 ISETDLYIGYRATSTDMTGTYYVFNADTKERTLSTDLTEGVVI 987

## RESULT 12

H84057  
 Pullulanase amyx [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 01-Dec-2000  
 C:Accession: H84057  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20263314  
 A:Accession: H84057  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-717 <STO>  
 A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA06983.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: amyx

Query Match 10.5%; Score 593.5; DB 2; Length 717;  
 Best Local Similarity 26.1%; Pred. No. 3.9e-28;  
 Matches 203; Conservative 99; Mismatches 290; Indels 187; Gaps 28;

QY 280 LISATQVQTAVIDDAYEAAEALSYGQNLADGVTFFVMAPTAQOVVYVYSAKKYIG 339  
 Db 89 LITVGRVVRTEAFDOQVYNGE--LGAlySHERTVFAVMPVATEVSLLEYENHHTLEA 145  
 QY 340 SHPMTRODSAGAWSOGSDLKGAFFRYAMTVYHPOSKRVEQYEVTDPAHSLSTNSEYS 399  
 Db 146 IEYQMKRSPNVMWSYSLDGLHGYCYLYRAVNGMKKTV-----DPAKSVTVNGENG 199  
 QY 400 QVVDLNDLSALPDGMDNLTMPHAOKTKADLAKMTHESHIRDSAMDOTVPAELRGKYIA 459  
 Db 200 VIFDLKRT--DPASMPVTHPIPEHT-----IYELHIRDAFYPGS--GHKKKGYLS 249  
 QY 460 LTA-----GDSNMVQHLKTLTASAGTVHVELLPVFDLATVNEFSDKVDADIQPSRLCE 512  
 Db 250 LABONTGCPMGWSTGLSTIREL--GVTHLELLPVADYGSDE----- 289  
 QY 513 VNSAVKSEFAGYCDGSGTVEEVLNQLKSDSDNPVOALNTLVAQTDSYNMGYDPFIY 572  
 Db 290 -----ANREKE-----YMGYDVVHY 305  
 QY 573 TVPEGSYATDP--EGTTRIKERTMTQAIKODIGMNVINDVYVNHNTMAAGPTDRTVLDKI 631  
 Db 306 FAPESYASDDEPDGVTYRKELKQLI--AOCORMDLRLVILDVYFNHYVM-----KESPFEXL 360  
 QY 632 VPMYQRLNETTGVSVEATCCSDSAPBEHRFAKLITADSLAVTWTYKIDGFFEDMGYAP 691  
 Db 361 VPGYIFRIDGEGNLADGTGCVNDIASERLMAVKKFTVDCVTVWASEYRVYDGFDFDLGIIID 420

QY 692 KAOLISAMERIKALNPDIYFFEGEWD-----SNQDRFEIASQINLXGTGIGFESDRLR 745  
 Db 421 QMTMEVRAALAPIRIROLILLGEGWDLTEPLENDKKTTLAQARAAPF-----LHFHDFR 476  
 QY 746 DSVKGGFPDSDGLARNOGIGSGAGVLPNELASLSDQVRHLADI,TRLCMAGNLADFWM 805  
 Db 477 DAVK--GSTFD-----LNKG--GFCAGNI-----ELKETLILASMGHEAQC-- 513  
 QY 806 IDKGAAGKSEIDYNAGAPGGAADPTEVYVYKSHDNQTLMDMSTYKASQADLATRR 865  
 Db 514 -DR-----PSQAIQYVEAHDMHTLMDKLOFPVPPDEEEVLR-- 549  
 QY 866 MQAVSLATVMLGCIAPDOGSELLRSKSFTRDSYSGDMFNRVDSLO-----DNNVAV 920  
 Db 550 CHRLATVILVTSQGIPLHGGEFFRTRKHGEGNSYNSPDWINGDMSRAITVAADEVYK 609  
 QY 921 GMPRIISDGSNY-----EVITRVKEMVATPGEAEIKOMTAFYQELTELKSSPLETLGD 974  
 Db 610 GILALRKARAFCEFTSFSEVNTNMCPL--KTPSEV-----IGFHYHHIAKCGPDEENVVLHN 664  
 QY 975 GSAVVKRYDERNTGSDQAAGLLVMTVDDGMAKAG-----ASLDSRLDGLVYAIN 1022  
 Db 665 GGG--DKVDIRLPASGEWE-----VIVDKKAGLYPLTFNGHTVSVOSRSSMVLRLN 716

## RESULT 13

C83375  
 Probable glycosyl hydrolase PA2160 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83375  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
 Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lachbig, K.; L  
 Norey, S.; Olson, M.V  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: C83375  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-716 <STO>  
 A:Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AA05548.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2160  
 C:Superfamily: glyx protein

Query Match 6.3%; Score 352.5; DB 2; Length 716;  
 Best Local Similarity 22.6%; Pred. No. 1.6e-13;  
 Matches 196; Conservative 114; Mismatches 286; Indels 273; Gaps 41;

QY 311 DG-GVTFKRYMAPTAQOVVYVYSAKKYIGSHPMTR-----DSASGAWWSOGSDLKGAFF 365  
 Db 25 DGLGVNFALFSAHAKVELCLDAR---GEKEIERIELPEVTDLMHGYLPDARHGQIY 80  
 QY 366 RYAM--TVYHPOS--RKVEQYEVTDPAHSLSTNSESO-----YVDLNDLSALKPDGMDL 417  
 Db 81 GYRVAGPIYEPDQGHENPNKLLIDPYAKOLVGRKXSEALFGYTTIGSADADLSFDRD-- 138  
 QY 418 TMPHAOKTKA-DIA-----KMTTHESHIRDSAMDOTVPAELRGKYIALTA 462  
 Db 139 SAPFVPSKVIDIPAFVMAERPVVRVPMRTVYEAHLRLSLSRHPOVPAVVGTFAGLNN 198  
 QY 463 GDSNMVQHLKTLTASAGTVHVELLPVFDLATVNEFSDKVDADIQPSRLCEVNSAVKSEF 522  
 Db 199 AD--LLAHLIRL---GVYSVELLPIT----- 219  
 QY 523 AGYCGSGTVEEVLNQLKSDSDNPVOALNTLVAQTDSYNMGYDPFIYTVPEGSYATD 582  
 Db 220 -GFVDDKHLLENGM-----SNWGINSTIAFFAPHPAYL-- 251





A:Cross-references: GB:AE001888; GB:AE000513; NID:96457936; PIDN:AAF09648.1; PID:9645793  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0264  
A:Map position: 1  
C:Superfamily: glyx protein

Query Match 6.1%; Score 344.5; DB 2; Length 720;  
Best Local Similarity 21.8%; Pred. No. 5e-13;  
Matches 198; Conservative 113; Mismatches 270; Indels 327; Gaps 45;

```
OY 306 GAQLADGVTFRVAPTAQOVVYVSADKKVISH-----PMTRDSASGASMSWGGSGLK 361
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 GATWDGKGTNFALYSENATGVELCLPDAE-----GHETRFPLEQT--AFVHGYLEPGI 68

OY 362 GAFRYAMVY---YHPOS--RKVEQEVTDPAHSLSTNSEYSOVVDLNDALSKPDGMDN 416
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 QPGORIGRYHGEYAPKGLRFNPVYLLDPYAKALDGTEDQDRGV---FGYVAGGEDD 124

OY 417 LTMPAOKTRADLA-----KMTIHESHIRDLASMDQTVPAELRGK 456
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 SQMGEEFQRGAPLGLVYDPMFNWGDQKPGIPEHQSVIYEAHVAGLITHTHPDVEBELRGT 184

OY 457 YLALTAGDSNMVQHLKTLASGVTHVELLPYEDLATVNEFSKVDADIQQPFSRLCEVNSA 516
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 YAGVAT--PALIDYLRDL--GITAIEFLPVHQ-----HVDDPF----- 218

OY 517 VKSEFAGYCDSCSTVEEVLNQLKQSDSQDNPOVALNTLVAQTDSYNMGYDPFHITYPE 576
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 -----LIDKGLT-----NY-WGYSTLNFFAPD 239

OY 577 GSYATD-----PEGTTRIKERTMLOAIKODLGNVIMDVYNNH---NAAGPTDRTSVL 628
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 VRYSAEARKGNPSGA--VPEKKNVRL- HDAGLEVILDVYNNHTAECHNMGPMSFKGI 296

OY 629 DKYVPMYQRLNET-----TGSVESATCCSDSAPEHRMFAKLADSLAVTTDYKIDG 681
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 297 DN--PTYRYLVADDOREFYFDYTGNSLNV-----RHPQTLQIMDSLRYWVTEMHVDC 348

OY 682 FRPDL-----MGYHPKQILISAMERIKALNP---DIYFEGGMSNOSDRELIASQIHLK 733
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 FRPDLASTLARGLEHVDQ--LSGFTTTHODPIISOVKLIAEPMDVGE----- 395

OY 734 GTGIGTF-----SDRLRDSVRCGCPFDSDALRONQIGSAGVLPNELASLSDQVR 786
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 396 GYQYGNFPMWMAEKNGTYYRDMR-----SFMKGSGGLASEIGY---RIT 436

OY 787 HLADELRLGMAGNLADPYMIDKGAAKKGEIDYNGAPGYAADPTEVYVYVSKHDNQL 846
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 GSSDL-----YEFNG-----RKPYASINFTVAHDGFTL 464

OY 847 WDMISY-----KASQEAD-----LATRVRMQAVS--LATVMILO 878
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 465 RDSVYTYEOKHNEANGEGNCHNINITWNCVEGPTDPEINRLRGQOMRNFATLLILGQ 524

OY 879 GIAFDQGSSELLRSKSFTRDSYSGDMFNRYVDSLQDNNYVGMPRISDDCSNEYEVITRV 938
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 525 GTPMILGDEFGRTGGGNNNAV-----CODN-----DISWYD----- 556

OY 939 KEMVATPGEALKQMTAFYOELTELKSSP-----LFTLG-----DGS 977
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 557 -----WEKYDEFELAFTRKILALRKAHPSLHRRKFPAGRNIRGEDVYDIWMLRFG-A 608

OY 978 VMKRVDFRNIGSDQAGLLVTVDDGMAKASLSR-----LDGLVYAINAA----- 1024
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 609 EMSDEDMNN--POTOSGMFLAGDGL--ADVAEGKPLTDDHLILLSSSYVDLPFKM 662

OY 1025 PESRTINEFAGETLQLSAIQOTAGENSILANGVQIAADGTVTLLPMSVAV--LELPQGEAQ 1082
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 663 PDLGGCGEM--DLLIDTSDGAE-----KYAAGGETTLRGRSVKLYRCQVPERKELE 712

OY 1083 GAGLPVSS 1090
```

Db 713 KLAEPAEA 720

Search completed: April 22, 2001, 09:04:57  
Job time: 4138 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 09:11:53 ; Search time 52.01 Seconds

(without alignments)  
718,568 Million cell updates/sec

Title: US-09-262-126c-4

Perfect score: 5633  
Sequence: 1 MLRYTRNALVLSLVLLSGC.....AVLELPGEAGAGLEPVSSK 1091

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5618.5	99.7	1090	1	PUL_A_KLEPN
2	4778.5	84.8	1096	1	PUL_A_KLEAE
3	703.5	12.5	843	1	PUL_A_THENA
4	352.5	6.3	721	1	GLGX_MYCTU
5	352.5	6.3	777	1	ISOA_FIASP
6	334	5.9	776	1	ISOA_PSEAY
7	334	5.9	776	1	ISOA_PSESP
8	302.5	5.4	657	1	GLGX_ECOLI
9	264	4.7	659	1	GLGX_HAETN
10	176	3.1	2334	1	MAPA_BACSU
11	169	3.0	580	1	YF62_MYCTU
12	167.5	3.0	666	1	GLGB_BACCL
13	166.5	3.0	731	1	GLGB_MYCTU
14	163	2.9	730	1	GLGB_HAETN
15	160.5	2.8	1861	1	APU_THETU
16	158.5	2.8	1289	1	VG34_BP74
17	156	2.8	1902	1	PIP_IACIC
18	152	2.7	728	1	GLGB_ECOLI
19	152	2.7	1902	1	P2P_IACIA
20	149.5	2.7	764	1	GLGB_STRAU
21	148.5	2.6	1090	1	GLGX_CELFI
22	148	2.6	1279	1	APU_THESA
23	147	2.6	627	1	GLGB_BACSU
24	147	2.6	940	1	SECA_STRGR
25	147	2.6	2366	1	TOXB_CLODI
26	146.5	2.6	995	1	ACAA_LTB57
27	146	2.6	1392	1	RPOB_NEIMB
28	145	2.6	1392	1	RPOB_NEIMA
29	143	2.5	1300	1	V338_MYCPN
30	143	2.5	1592	1	GTF2_STRDO
31	141.5	2.5	880	1	LYTD_BACSU
32	140	2.5	1597	1	GTF1_STRDO
33	140	2.5	1902	1	P2P_LACPA

34	139	2.5	498	1	AMY3_DICTH
35	139	2.5	520	1	AMY_BACME
36	139	2.5	2292	1	POLG_EMCVB
37	139	2.5	2292	1	POLG_EMCVD
38	137.5	2.4	773	1	GLGB_SYN7
39	136.5	2.4	591	1	CDAS_BACSH
40	136	2.4	639	1	GLGB_BACST
41	136	2.4	713	1	AMYR_BACS8
42	135.5	2.4	861	1	GLGB_SOLRU
43	135.5	2.4	1902	1	PIP_IACIC
44	135	2.4	639	1	GLGB_BUTFI
45	135	2.4	919	1	AMY_STRLI

## ALIGNMENTS

```

RESULT 1
PUL_A_KLEPN STANDARD: PRT: 1090 AA.
AC P07206:
DT 01-APR-1998 (Rel. 07, Created)
DT 01-APR-1999 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PULULANASE PRECURSOR (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-ALPHA-
DE GLUCOSIDASE) (PULULAN 6-GLUCANOHYDROLASE).
GN PUL_A.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UNE 5023;
RX MEDLINE=90205629; PubMed=2181242;
RA Kornacker M.G., Pugsley A.P.;
RT Molecular characterization of pulA and its product, pullulanase, a
RT secreted enzyme of Klebsiella pneumoniae UNF5023.*;
RL Mol. Microbiol. 4:73-85(1990).
[2]
RN [3]
RP SEQUENCE OF 944-1090 FROM N.A.
RX MEDLINE=86033621; PubMed=3902792;
RA Chapon C., Raibaud O.;
RT *Structure of two divergent promoters located in front of the gene
RT encoding pullulanase in Klebsiella pneumoniae and positively
RT regulated by the malt product.*;
RL J. Bacteriol. 164:639-645(1985).
[3]
RN [3]
RP SEQUENCE OF 944-1090 FROM N.A.
RX MEDLINE=89291709; PubMed=2661532;
RA D'Enfert C., Pugsley A.P.;
RT *Klebsiella pneumoniae pulS gene encodes an outer membrane
RT lipoprotein required for pullulanase secretion.*;
RL J. Bacteriol. 171:3673-3679(1989).
-1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYSES
(1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO
FORM MALTOSE.
-1- SUBUNIT: HOMOTRIMER.
-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
ANCHOR (PROBABLE).
-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL: X52181; CAA36431.1; -
EMBL: M12503; AAA25087.2; -

```

DR EMBL: M29097; AAA61976.1; -  
 DR PIR: A25025; A25025.  
 DR PIR: A32880; A32880.  
 DR PIR: S11823; S11823.  
 DR InterPro: IPR000461; -  
 DR Pfam: PF00128; alpha-amylase; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hydrolyase; Glycosidase; Membrane; Lipoprotein; signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1090  
 FT LIPID 20 20  
 FT ACT\_SITE 684 684  
 FT ACT\_SITE 713 713  
 FT ACT\_SITE 841 841  
 FT CONFLICT 6 6  
 FT CONFLICT 10 10  
 FT CONFLICT 15 15  
 FT CONFLICT 23 23  
 FT CONFLICT 31 31  
 FT CONFLICT 34 34  
 FT CONFLICT 36 36  
 FT CONFLICT 55 58  
 FT CONFLICT 58 58  
 SO SEQUENCE 1090 AA; 118098 MM; 240AETDFB3PFIIBD6 CRC64;

Query Match 99.7%; Score 5618.5; DB 1; Length 1090;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1090; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLRTTRNALVGLSVLLSGDNGSSSSGPNPTPDNDVYVRLPVAVPGEAVTVEN 60  
 DB 1 MLRTTRNALVGLSVLLSGDNGSSSSGPNPTPDNDVYVRLPVAVPGEAVTVEN 59  
 QY 61 QAVIHLVDIAGITSSSADYSSKNLYLMNNETCDALSAVADNDVSTTPSGSDKGPYV 120  
 DB 60 QAVIHLVDIAGITSSSADYSSKNLYLMNNETCDALSAVADNDVSTTPSGSDKGPYV 119  
 QY 121 VIPLANKSGGCIYVIRGTGKLLDSDLRVAFGDTDTVSYIAGNSAVYDSRADAFRAAF 180  
 DB 120 VIPLANKSGGCIYVIRGTGKLLDSDLRVAFGDTDTVSYIAGNSAVYDSRADAFRAAF 179  
 QY 181 GVALAEAMVDKNTLLMPGODKPIYRLYYSIHSKVAADGEGKTDYRLKLTPTTVSQV 240  
 DB 180 GVALAEAMVDKNTLLMPGODKPIYRLYYSIHSKVAADGEGKTDYRLKLTPTTVSQV 239  
 QY 241 SMRPPLSSYAAFKLPDANVDELLOGETVAIAAEDGILISATOVOTAGVLDVAVEAA 300  
 DB 240 SMRPPLSSYAAFKLPDANVDELLOGETVAIAAEDGILISATOVOTAGVLDVAVEAA 299  
 QY 301 EALSYGQNLADGVTFFRVNAPTAQOVVVVYVYSDKVIYGHSHPMTRDSASGMSWQGGSD 360  
 DB 300 EALSYGQNLADGVTFFRVNAPTAQOVVVVYVYSDKVIYGHSHPMTRDSASGMSWQGGSD 359  
 QY 361 KGAFYRAMTVIIPQSRKVEQYETDPYAHSLSTNSEYSGVVDLNDALKPDGDNLTMP 420  
 DB 360 KGAFYRAMTVIIPQSRKVEQYETDPYAHSLSTNSEYSGVVDLNDALKPDGDNLTMP 419  
 QY 421 HAQTKADLAKMTIHESHIRDLSAMDTVPABELRCKYIALTAGDSNNVQHLKTLASAGVT 480  
 DB 420 HAQTKADLAKMTIHESHIRDLSAMDTVPABELRCKYIALTAGDSNNVQHLKTLASAGVT 479  
 QY 481 IVELLPEDLATVNEFSDKVAADIQOPSRICEVNSAVKSSSEFAGCDSGSTEVEVLNOLK 540  
 DB 480 IVELLPEDLATVNEFSDKVAADIQOPSRICEVNSAVKSSSEFAGCDSGSTEVEVLNOLK 539  
 QY 541 QSDSDNPQVQALNTLVAGTDSYNMGYDPFHYYTPEGSYATDPEGTRIKERFMIAQIK 600  
 DB 540 QSDSDNPQVQALNTLVAGTDSYNMGYDPFHYYTPEGSYATDPEGTRIKERFMIAQIK 599  
 QY 601 QDLGMNVIMDVVYVYHNTAAGPTDRTSVLDKIVPVYQYRLNETTGSVESATCCSDSABEHR 660  
 DB 600 QDLGMNVIMDVVYVYHNTAAGPTDRTSVLDKIVPVYQYRLNETTGSVESATCCSDSABEHR 659

QY 661 MPAKLIADSLAVTTDYKIDGFRFDLMGYHPKAOILSAMERIKALNDPIYFEGGMSNQ 720  
 DB 660 MPAKLIADSLAVTTDYKIDGFRFDLMGYHPKAOILSAMERIKALNDPIYFEGGMSNQ 719  
 QY 721 SDRFELASQINLKGCTGCTPSDRLRDSVRGGCPDSDGDLARONCIGSGAGVLPNELASL 780  
 DB 720 SDRFELASQINLKGCTGCTPSDRLRDSVRGGCPDSDGDLARONCIGSGAGVLPNELASL 779  
 QY 781 SDDQVRHLADLTRIGMAGNLADFPWIDKCAAKKGSGLDVGAPGAAAPTEEVANVYSK 840  
 DB 780 SDDQVRHLADLTRIGMAGNLADFPWIDKCAAKKGSGLDVGAPGAAAPTEEVANVYSK 839  
 QY 841 HDNQLTMDMISYKASQADLATRVKQAVSLATVYLCOGIAFDQGSSELLRSKSTFDSY 900  
 DB 840 HDNQLTMDMISYKASQADLATRVKQAVSLATVYLCOGIAFDQGSSELLRSKSTFDSY 899  
 QY 901 DSGMFRVYVSLQDNNYVNGMPRISSDGSNVEYITRVKEMVATPGAEIKOMTAFYQEL 960  
 DB 900 DSGMFRVYVSLQDNNYVNGMPRISSDGSNVEYITRVKEMVATPGAEIKOMTAFYQEL 959  
 QY 961 TELKSSPLFTIGDSAVMKRVDFRNTGSDQAGLLVWTVDDGKAKASLDSRLDGLVVA 1020  
 DB 960 TELKSSPLFTIGDSAVMKRVDFRNTGSDQAGLLVWTVDDGKAKASLDSRLDGLVVA 1019  
 QY 1021 INAPESRTINEFAGETLQLSAIOCTAGENSLANGVOIADGVTLPAMSVAVLELPQGE 1080  
 DB 1020 INAPESRTINEFAGETLQLSAIOCTAGENSLANGVOIADGVTLPAMSVAVLELPQGE 1079  
 QY 1081 AOCAGLPVSSK 1091  
 DB 1080 AOCAGLPVSSK 1090

RESULT 2  
 PUA KLEAE  
 ID PUA KLEAE STANDARD; PRT; 1096 AA.  
 AC P07811;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE PULIDANASE PRECURSOR (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-ALPHA-DE GLUCOSIDASE) (PULULAN 6-GLUCANOHYDROLASE).  
 GN PUA.  
 OS Klebsiella aerogenes.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxID=28451;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H70;  
 RX MEDLINE=67194626; PubMed=3155373;  
 RA Katsuragi N., Takizawa N., Murooka Y.;  
 RT "Entire nucleotide sequence of the pullulanase gene of Klebsiella aerogenes H70".  
 RT J. Bacteriol. 169:2301-2306(1987).  
 CC -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLASES (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO FORM MALTOSE.  
 CC -1- SUBUNIT: HOMOTRIMER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID ANCHOR (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M16187; AAA25124.1; ALT\_SEQ.

```

DR   PIR: A26879. A26879.
DR   InterPro: IPR000461.
DR   Pfam: PF00128; alpha-amylase; 1.
DR   PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Hydrolyase; Glycosidase; Membrane; Lipoprotein; signal.
FT   SIGNAL      1      19
FT   CHAIN       1      196
FT   LIPID       20      20
FT   ACT_SITE    694      694
FT   ACT_SITE    723      723
FT   ACT_SITE    851      851
SQ   SEQUENCE    1096 AA; 119335 MW; FE7D9167CDACED79 CRC64;

Query Match      84.8%; Score 4778.5; DB 1; Length 1096;
Best Local Similarity 84.5%; Pred. No. 1.6e-261;
Matches 934; Conservative 49; Mismatches 77; Indels 45; Gaps 9;

QY   1 MLRTIRNALVIGSLVLLSGCDNGSSSSSSSS-----GNPDTPDNDVVRLLPDVA 49
DB   1 MLRYCHALFLGSLVLLSGCDNSSSSSSSGSPGPNPCNPCTPGTDPDQVVRLLPDVA 60

QY   50 VPGEAVTAVEAQAIVHLVDIAGITSSSAADYSKMLYLMMNTCCALSPADMDVSTT 109
DB   61 VPGEAVQAASQAQVHLVDIAGITSSPADYATKMLYLMMNETCALSPADMDVSTT 120

QY   110 PGSDKPYPIVWIPILNKSQGINVIRGDTDKLIDSLRFAFGDPTDRVSIAGNSAVY 169
DB   121 PFGSKPYPIVWIPILNKSQGINVIRGDTDKLIDSG-RVSTSDPTDRVSIAGNSAVY 179

QY   170 DSRADAFRAAFGVALAEAHWDKNTLLMPGQDKPIVRLYSSHSSKVAADGKFTDRL 229
DB   180 DSRADAFRAAFGVALADAHWDKNTLLMPGGENKPIVRLYSSHSSKVAADNGEFDKYV 239

QY   230 KLPTPTVSOQVSMRPHLSSTAAFLKPNANVDLLOGETVAIAADGCLL-----SAQ 285
DB   240 KLPTPTVNOQVSMRPHLSSTAAFLKPDVNDVDELQDGGGIAES-DEILSLSPGADR 298

QY   286 VQTAGVLDAAVEAAEALSYGAQLADGVTFEVAAPTAQOVVVVYSADKKYIGSHPMTR 345
DB   299 RRAGRYL-----CRRAEALSYGAQLADGVTFEVAAPTAQOVVVVYSADKKYIAHSHPMTR 354

QY   346 DSASGAMWGGSDLKGAFTYAMTVYHPQSRKVEQYEYTDPAHSLTNSSEYSOVDLNL 405
DB   355 DSASGAMWGGSDLKGAFTYAMTVYHPQSRKVEQYEYTDPAHSLTNSSEYSOVDLNL 414

QY   406 DSAALKPDGMDNLTMHAKTKADAKMTIHESHIRDLMSMDQTVPAELRGKYLALTAGDS 465
DB   415 DSAALKPEBGMDGLTMHAKTKADAKMTIHESHIRDLMSMDQTVPAELRGKYLALTAGDS 474

QY   466 NMVOHLKRLSASGVTHVELLPVFDLATVNEFSDKVADIQQPSRLCEVNSAVKSSFPAGY 525
DB   475 NMVOHLKRLSASGVTHIELLPVFDLATVNEFSDKVADIQQPSRLCEVNSAVKSSFPAGY 534

QY   536 CDSSGSTEVEVLNQLKQSDSDQNPQVQALNTLVAQTDSTYNNMGYPHRYTVPEGSYATDPEG 585
DB   535 CDSSGSTEVEVLNQLKQSDSDQNPQVQALNTLVAQTDSTYNNMGYPHRYTVPEGSYATDPEG 594

QY   586 TTRIKFEPTMIOAKQDGLGMNVYVNNHTNAAPPTDRTSYLKLIVRYRYRLMETTGS 645
DB   595 TARIKEPTMIOAKQDGLGMNVYVNNHTNAAPPTDRTSYLKLIVRYRYRLMETTGS 654

QY   646 VESATCCSDSAPERRHMFAKLIADSLAVVTTDYKIDGFRFDLMGYAPKAOIISAMERIKAL 705
DB   655 VESATCCSDSAPERRHMFAKLIADSLAVVTTDYKIDGFRFDLMGYAPKAOIISAMERIKAL 714

QY   706 NPDIYFEGEGMDNSQDRFEELASQINLKGITGTFSSRLRDSVRGGGPFDSGDALRQNG 765
DB   715 NPDIYFEGEGMDNSQDRFEELASQINLKGITGTFSSRLRDAVRGGGPFDSGDALRQNG 774

QY   766 IGSAGCVLPNELIASLSDQVVRHLADLTRLGMAGNLADVFMIDKDAAKGSIIDYNGARG 825
DB   775 VGSAGCVLPNELITLSDQVVRHLADLTRLGMAGNLADVFLIDKDAVVRGSEIDYNGARG 834

```

```

QY   826 GYAADPTEVVNVYSKHDNLTMDMISTYKASQEAADLATRVMAQVSLATVMLGQIAFDQ 885
DB   835 GYAADPTEVVNVYSKHDNLTMDMISTYKASQEAADLATRVMAQVSLATVMLGQIAFDQ 894

QY   886 GSELLRSKSFSTDSDSGDFNRVDYSLQDNNYNGMPRIISDGSNVEYTRKEVATP 945
DB   895 GSELLRSKSFSTDSDSGDFNRVDYSLQDNNYNGMPRIISDGSNVEYTRKEVATP 954

QY   946 GSAELKOMAFYOELETLLKSSPLFLTIGSAVMKRVDFRNTGSDQAGLLVMTVDGK 1005
DB   955 GETELKOMAFYOELETLLKSSPLFLTIGSAVMKRVDFRNTGSDQAGLLVMTVDGK 1014

QY   1006 AGASL-----DSRLDGLVVAIINAPESRTLNEFAGETLQLSA--IQOTAGENSIA 1053
DB   1015 AGRSGQGPCRRRRRGQGRAG-----KPDAAGLRRIAPARVYSAGGGRVYAGER--- 1064

QY   1054 NGVQIADGTVTLPAMSAVLELPQ 1078
DB   1065 --VQVAADGSVTLPAMSAVLELPQ 1087

RESULT      3
PULA_THEME  STANDARD:      PRT:      843 AA.
ID           033840;
AC           30-MAY-2000 (Rel. 39, Created)
DT           30-MAY-2000 (Rel. 39, Last sequence update)
DI           30-MAY-2000 (Rel. 39, Last annotation update)
DE           PULLULANASE PRECURSOR (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-ALPHA-
GN           GLUCOSIDASE) (PULLULAN 6-GLUCANOHYDROLASE).
OS           PULA OR TM1845.
OC           Thermotoga maritima.
OX           Bacteria; Thermotogales; Thermotoga.
NCBI_Taxid=23336;
RN           [1]
RP           SEQUENCE FROM N.A.
RC           STRAIN=MSB8 / DSM 3109;
RA           MEDLINE=99287316; PubMed=10360571;
RA           Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA           Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA           McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA           Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA           Selberg J., Sutton G.C., Fleischmann R.D., Eisen J.A., White O.,
RA           Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT           "Evidence for lateral gene transfer between Archaea and Bacteria from
RT           genome sequence of Thermotoga maritima.";
RL           Nature 399:323-329(1999).
CC           -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYSES
CC           (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO
CC           FORM MALTOSE.
CC           -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC           KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC           This SWISS-PROT entry is copyright. It is produced through a collaboration
CC           between the Swiss Institute of Bioinformatics and the EMBL outstation
CC           the European Bioinformatics Institute. There are no restrictions on its
CC           use by non-profit institutions as long as its content is in no way
CC           modified and this statement is not removed. Usage by and for commercial
CC           entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC           or send an email to license@sib-sib.ch).
CC           EMBL: AJ001087; CA040522.1; -.
CC           DR           EMBL: AE001821; AAD36907.1; -.
CC           TIGR: TM1845; -.

```



QY	362	GAFFRATVATYHP-----OSRKVEQYVTPDYPAHSLSTNSEYQVY--DLDNSALAKPGP	413
Db	80	TPGQRIGERRVINGPDPAPGAHRCRCPSTLLDPYKSGKHGDDTFEQALYSTVDN--AYDPUS	137
QY	414	W-----DNL--TW-----PHAQKTKADLAKMTIHESHIRIDLSAMDQVTPAELRCK	456
Db	138	TPPMVDSLGHTMTSVINPFEDMAADYDSPRTPEHYVIEAHVKGMTQHPISLPELRGT	197
QY	457	YLTALTAGDSNNVQHLKTLSSAGYTHVLLVYFDLAYNEFSKDVADIQQPFSRLCEVNSA	516
Db	198	YAGL--AHPIYLIIDLNELN--YVAVELMP-----VHOFLMD-----SKLIDLG--	236
QY	517	VKSEFAQYCCSGSTVEEVNLQOLKOSDQONQOVQALNTFLAQTDSTYNGMDPEFHYTPE	576
Db	237	-----LKNY-----WQNTFGGFAPH	252
QY	577	GSVATDPDEGTRRIKEFRTMIOAIKODLGMYIMDVVYNYHT--NAAGPDRTSVLDKIVP	633
Db	253	HQVASTROAGSAVAEFTPTVYRST--HEAGIEVILIDVYVNYHTAEGNHLIGTINFRGIDNTA-	310
QY	634	MYQRLNETTGSVESATCCSDSA--PEIRMTAKLIADSLAVMTDVIYIDGFRDLNGYHPR	692
Db	311	YVRLMHDLREFYDDFTGTSNLSNARPHPTLQILIMDSLRVYTEMVADDFRDLAS--TL	367
QY	693	AQIISAMERIKALNPDIYFFGEGWDSNQSOFELIASOINLKGTGIGTFSDRILDSVRCGG	752
Db	368	ARELHVDRLSLA-----FF-----DLVQDD--FVYSQVKLIAE-----	398
QY	753	PPDSGDLAKNQGIGSGAGVLPNELASLSDQYVRHLADLTRIGMAGNLADFPVIMDKGAA	812
Db	399	PMVDGEGGYQ---VGNPPG--LMTFWMNGKYRDYVR--DYWR--CEPATLGEF-----AS	443
QY	813	KKSGEIDYNGAPGTYADPFEVNNYVSKHDNQLTMDMISY--KASQDAD-----	859
Db	444	RLTGSSDLYEATG--RRPASINIEFYAHGFTLMDLYSTNDKHNANEENNRNDESYNR	500
QY	860	-----LATRYMQAVSLATYMWLGGLAFPOQSGSEILRSKSFTRDSYSGD	904
Db	501	SMWNGVGEPTDDPDLALRRQRKMRNMATIMWSQGTPTMHAHGEIGRTGYGNNNYCQDS	560
QY	905	WPNRVDYSLQDNNNYNVMGPRIISDGSNYEIVTRVKEKMATPGEAEIKQMTAFYQELTELR	964
Db	561	ELSMWMDSLVADKN-----ADLLAFARAKATYLR	587
QY	965	KSSPLF---TLGDSASAMKRVDFRN-----TGSQOQAGLWYVYDGSMAKGAISDSRLGL	1011
Db	588	KNHKVFRRRRFESEPIRSGDEVRDIAWLTPSGRE--MHEHW--GRGFDR-----	634
QY	1018	VVAINAAPESTRLEAFEGFTIQLSLAI-----OQTAGENSLANGY	1056
Db	635	CVAVFLNGEATITAPDAGSERVDDSFLLCENADHDHVEYFVPHDGYAQQWGTGELTDINVEY	694
QY	1057	-----QIAADGVTVLPAMSYAVL	1074
Db	695	GDIDLTFTATDTFTSVPARSLVL	717
RESULT 5			
ISOA_FLASP	ISOA_FLASP	STANDARD;	PRT; 777 AA.
AC	032611;		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	15-DEC-1998 (Rel. 37, Last annotation update)		
DE	ISOAMT_LASE PRECURSOR (EC 3.2.1.68).		
GN	IAM.		
OS	Flavobacterium sp.		
OC	Bacteria; CFB group; Flavobacteriia; Flavobacteriaceae; Flavobacterium		
OX	NCBI_TaxId:239;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-97340931;		
RX	pubmed-9197405;		

RA	Kirohn B.M., Barry G.F., Kishore G.M.:	
RT	"An isoamylase with neutral pH optimum from a <i>Flavobacterium</i> species:	
RL	cloning, characterization and expression of the <i>lam</i> gene."	
RI	Mol. Gen. Genet. 254:469-478(1997).	
CC	-1- FUNCTION: HAS A HIGH RATE OF HYDROLYSIS FOR GLYCOGEN. DOES NOT	
CC	CLEAVE PULIDAN. HAS A PH OPTIMUM OF 6-7.	
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH	
CC	LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.	
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO	
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; U90120; AAB6356.1; -.	
DR	InterPro; IPR000461; -.	
DR	Pfam; PF00128; alpha-amylase; 1.	
DR	KW Hydrolyase; Glycosidase; Signal.	
FT	SIGNAL	1 32
FT	CHAIN	33 777
FT	ACT_SITE	410 410
FT	ACT_SITE	478 478
FT	ACT_SITE	533 533
FT	DISULFID	419 423
FT	SEQUENCE	777 AA; 84340 MW; 90548988BFD1445B CRC64;
FT		POTENTIAL.
FT		ISOAMYLASE.
FT		BY SIMILARITY.
FT		BY SIMILARITY.
FT		BY SIMILARITY.
FT		BY SIMILARITY.

[illegible]

```
Db 472 FPGMAEMWNLKDALKKKMLGCVETVDTGTLA-----TR--PAGS----- 511
QY 805 MIDKGAARKGSEIDYNGAPGYAADPTEVYVNSKHNDOTLMDISYKASQEND----- 859
Db 512 -----NDLVCDDGCRKPMHSINTVVAHDGFTLMDIAYNKNQNNQPPRYG 555
QY 860 -----LATRVKMOAVSLATVVLGOCIAFDQOCSELRKSKFTROSYDS 902
Db 556 PSDGDEHNLMSWNOGGVVAQORAKARTGLALLMSAGVPMITGDEALRQ----- 606
QY 903 GDMFNRYDYSLODNVNVVWCPRISSDGSNYEVLTVKVENATPGSEALKQMPAFYOELTE 962
Db 607 ---FC-----NNNTYVL-----DSANMLYSR--SALEADHETTTKRLIAFRKAHPA 649
QY 963 LRKSSPLFTLGDSSAVKRYD-FRNTGSDQOAGLVTVDGKAKA---SLDSRLDGLV 1018
Db 650 LRANFYASADTNGNVMEQLRMFKPDCAQ-----AOSAVFNGADNMLALMRIDGSE 700
QY 1019 VAINAAPESRTLEPAGEF-----LQSLAQOTFACENSLANGVOIAADQVTL 1066
Db 701 FGDSASAIYVAWNGWGVGAVDEKLPWPGTKQWRYVTDATATNEGPNVAV--ALPGSETL 756

RESULT 6
ID ISOA_PSEAY STANDARD; PRT; 776 AA.
AC P10342;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-DEC-1998 (Rel. 37, last sequence update)
DE ISOAMYLAZE PRECURSOR (EC 3.2.1.68).
GN IAM.
OS Pseudomonas amylocleramosa.
OC Bacteria: Proteobacteria.
OX NCBI_TaxID=32043;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SB-15;
RA AMEMURA A., CHAKRABORTY R., FUJITA M., NOUMI T., FUTAI M.;
RT "Cloning and nucleotide sequence of the isoamylase gene from
Pseudomonas amylocleramosa SB-15."
RL J. Biol. Chem. 263:9271-9275(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JD210;
RA CHEN J.H., CHEN Z.Y., CHOW T.Y., CHEN J.C., TAN S.T., HSU W.H.;
RT "Nucleotide sequence and expression of the isoamylase gene from an
isoamylase-hyperproducing mutant, Pseudomonas amylocleramosa JD210."
RL Biolum. Biophys. Acta 1087:309-315(1990).
RN [3]
RP SEQUENCE OF 744-776 FROM N.A.
RC STRAIN-SB-15;
RA AMEMURA A., FUJITA M., FUTAI M.;
RT "Transcription of the isoamylase gene (iam) in Pseudomonas
amylocleramosa SB-15."
RL J. Bacteriol. 171:4320-4325(1989).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RA MEDLINE=98387895; PubMed=9719642;
RT "Three-dimensional structure of Pseudomonas isoamylase at 2.2-A
resolution."
RL J. Mol. Biol. 281:885-897(1998).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH
LINKAGES IN GLYCOGEN, AMILOPECTIN AND THEIR BETA-LIMITS DEXTRINS.
CC -1- INDUCTION: BY MALTOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLAZE SUBFAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: J03871; AAA25854.1; -
DR EMBL: X13378; CA31754.1; -
DR PIR: A28109; A28109.
DR PIR: 1BF2; 12-AUG-98.
DR InterPro: IPR000461; -
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 776
FT ACT_SITE 401 401
FT ACT_SITE 481 481
FT ACT_SITE 536 536
FT DISULFID 410 422
FT DISULFID 546 616
FT DISULFID 738 766
FT CONFLICT 8 8
FT CONFLICT 126 126
FT CONFLICT 169 171
FT CONFLICT 386 386
FT CONFLICT 413 416
FT CONFLICT 454 489
FT FT
FT CONFLICT 555 556
FT CONFLICT 650 657
FT SEQUENCE 776 AA; 83626 MW; F738BF8040246169 CRC64;

Query Match 5.98; Score 334; DB 1; Length 776;
Best Local Similarity 22.5%; Pred. No. 1,8e-11;
Matches 176; Conservative 86; Mismatches 242; Indels 278; Gaps 32;

QY 304 SYGQALDGGVTFVRVMAPTQAOVNVVYSADKKYIGHPMTRDGA-SGAN-----SMQG 356
Db 36 SYDAQOAN--ITFRVYSSQATRIVLVYLSAGYQVQESATTLTPAGSGVAIVVPVSSIK 93
QY 357 GSDKGAFFRYAMTVYHPOSRYVEYEVTPYAHSLSTNSEYSOV--VDLNSALRPDGM 414
Db 94 AAGITGAVY-YGRAMGPMN-----PYASNMCKGQAGFVSDVANGDRFNPN-- 140
QY 415 DNLTPHAKRTKAD-----LAKMT----- 433
Db 141 KLLDPRPAQEVSDPLPNSNONGNVEFASGASYRTTDSGIYAPKGVVLVPSTOSTGTRPTR 200
QY 434 -----IHESHIDLSAMDQTVPAELRGKYLALTAGSNMYOHLKT--LSASGVTHVEL 485
Db 201 AQRDQVYIEVHVGFTTPQDPSIPAQVRYG--AG-----LKASYLASLGTVAEFL 251
QY 486 PVFDLATVNEFSKVDADIQPFSSRLCEVNSAVKSEFAGYCDGSGIYEEVLNLDKQSDQ 545
Db 252 PVQEE--TQNDANDVVPN-----SDAN 270
QY 546 DNPQVQALNTLVAGQTSVNMNGYDPFHVTPPEGVAVDPRECTRIKEFRITQIAIKQDLGM 605
Db 271 QN-----YGVYMTENFSPDRRAVYKKAAGPTAEFOAVVQAF-HNAGI 313
QY 606 NVIMDVVYNNHTNAG-----PTDRTSVLDKIVPM-----YYORLNETTGSVESATCC 652
Db 314 KYVMDVYVNNHTAGCGTWTSSDPTTAT-----IYSWRQDLNATYVELLSQGYFYDNGIG 368
QY 653 SDSAPHEHMAFLADSLAVWTDYKIDGFRFDLMGYPHQAQLISAMERIKALNPDIYFF 712
Db 369 ANNTYVTVAQNILVDSLAVWANTMGVDGFRFDLASVLSGNSCLNGAVTASAPCPNGGYN 428
QY 713 GEGMDSNOSDRFELASQINK-----GNGIGTFSDRLR--DSVRGGCPDPDSC----- 757
```





```

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mao B., Shao Y.;
RT *The complete genome sequence of Escherichia coli K-12.*;
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE OF 1-590 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=89108020; PubMed=2975249;
RA Romeo T., Kumar A., Preles J.;
RT *Analysis of the Escherichia coli glycogen gene cluster suggests that
catabolic enzymes are encoded among the biosynthetic genes.*;
RL Gene 70:363-376(1988).
CC -1- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A STOP
CODON IN POSITION 507.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01616; AAA88735.1; ALT_SEQ.
DR EMBL: U18997; AAA58228.1; -.
DR EMBL: AE000419; AAC76456.1; -.
DR PIR: J10400; BVCEGX.
DR HSSP: P10342; 1BP2.
DR Ecogene: EG10381; g19X.
DR InterPro: IPR000461; -.
DR Pfam: PF00128; alpha-amylase; 1.
KW Glycogen biosynthesis; Hydrolase; Glycosidase.
FT ACT_SITE 336 BY SIMILARITY.
FT ACT_SITE 443 BY SIMILARITY.
FT CONFLICT 288 Y -> Y (IN REF. 2).
SQ SEQUENCE 657 AA; 73576 MW; AA9EFC1F67FD0420 CRC64;

```

Query Match 5.4%; Score 302.5; DB 1; Length 657;  
 Best Local Similarity 21.7%; Pred. No. 8.1e-10;  
 Matches 187; Conservative 104; Mismatches 266; Indels 303; Gaps 41;

```

DB 202 - -AQA-----SEPRLOKRL-----SNWGTNPVAFALHP 231
QY 578 SVATDEGSTRKEPRFMIOAKIDGMNIMDVYVNT--NAAGPTDTSVLIDIVP 634
DB 232 AYACSPB--TALDEFDAIKALHK-AGEVILDIVLNHSALDLDGLSLRGIDN-RST 287
QY 635 YQRLNETTGSVESATCCSDSAP-EHRMFAKLIADSLAVTDTYKIDGFRFDL--NGYH 690
DB 288 YWIR---EDGDYHNMTGCGTLMLSHPAVVDVASACLRVVERCHVDGFRFDLAAVWGRT 344
QY 691 PK---AQILSNERIKALNLPDIYFPEGGDSNQSDFELASINLKGIGTGF----- 740
DB 345 PEFRDAPLEFALQNCPLVS-QVKLIAEPWD-----IABGQYGVNFPPLFAE 391
QY 741 -SDRLDSYVGGPFDSDPALRONOGISGAGVLPNELASLSDQYRHLADLRILGMAGN 759
DB 392 WNHDFDARFRF-----LHIDLPGLAFAG---RFAASDVFRRN-----GR 430
QY 800 LADFYMIDKGAARKGSEIDYNGAPGYAADPTEVYVYSKHDNQTLMDSY-----KA 854
DB 431 L-----PSAINTLVTHDGFLLRDCVCFNRKHNEA 460
QY 855 SQE-----ADLAFVVR--WQAVSLATVMLGGIATFQGSSEL 890
DB 461 NGENRNDGTNNNTSNHKEGGLGSLDLVERRDSIHAL-LTLLLSQGTPLMADDEG 519
QY 891 RSKSFTRDSYDSDGMFNRVDYSLQNNYVNGMPRISDGSNVEITRVKEMVATPGAE 950
DB 520 HSHGNNMNV-----CODNQ-----LTVLDSQASG----- 546
QY 951 KQNTAFYQELTELKSSPL-----FTLGDGSAVMKRVDFRNTGSDQAGLVMTVDGM 1004
DB 547 - -LTAFTALILHKRIPALVERNWEEDGN-----VRLNRYAQP-----LSTDEM 593
QY 1005 KAKASLDSRL-DELVVAINAPESRTLNEFAGETQLSAIQOATAGENSILANGVIAADGT 1063
DB 594 NGPKQIILLSDRFLAINNTLVEITVLTPAGE--WHALPPAGE-----DNP 639
QY 1064 VTLPAVAVALELPGEAAG 1083
DB 640 VITAVW-----QGPAHG 651

```

RESULT 9  
 GLOX\_HAEIN STANDARD; PRT; 659 AA.  
 ID GLOX\_HAEIN  
 AC P45178;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DE GLYCOGEN OPERON PROTEIN GLOX (EC 3.2.1.1).  
 GN GLOX OR H11358.  
 OS Haemophilus proteobacteria.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fieschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 Ravelange A.R., Bull C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 Scott J.D., Shirley R., Liu L.-I., Glodde A., Kelley J.M.,  
 Meldman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 Uitterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
 Fine L.D., Fitchman J.L., Fuhmann J.L., Georgagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 Venter J.C.;  
 RT \*Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd.\*;  
 RL Science 269:496-512(1995).

CC -1- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/  
 CC CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U32815; AAC23005.1; -  
 CC DR HSSP: P10342; 1BF2.  
 CC DR TIGR: H11358; -  
 CC DR InterPro: IPR000461; -  
 CC Pfam: PF00128; alpha-amylase; 1.  
 CC Glycogen biosynthesis; Hydrolase; glycosidase.  
 CC ACT\_SITE 333 333 BY SIMILARITY.  
 CC ACT\_SITE 440 440 BY SIMILARITY.  
 CC SEQUENCE 659 AA; 75290 MW; E2B71F610E8B1C6C CRC64;

Query Match 4.7%; Score 264; DB 1; Length 659;  
 Best Local Similarity 22.0%; Pred. No. 1.2e-07;  
 Matches 166; Conservative 81; Mismatches 209; Indels 298; Gaps 39;

OY 367 YAMFY-----HPOSRRVQYEYTDPAHSLS-----TNSFSQ-----YVLDNSA-LKPRG 413  
 DB 74 YAFRIHGFANPDKL-----ILDPAKAVNGKPKDLSSESSSWFLSNRNNAHLAPRA 127  
 OY 414 -----MDLTPHAKTKADLAKMTIHESIRDLASMDQTPALRGKYATATAGDS 465  
 DB 128 VVISEEFEMENDTSPN-----TPMAETIYELHVGKFSQLNKIKIPALRGTYTGL-AHVP 181  
 OY 466 NKVOHLKTLASGVTHVELLPV-EDLATVYNEFSKVDADIQPFSLCEVNASVKSEFAG 524  
 DB 182 NLA-LKLEL-GYAVLELLPVNFH-INE----- 205  
 OY 525 YDSSGTYEVLNQLKQSDSQNPQVLAQTDSDNMGYDPFHYTPREGSVATDPE 584  
 DB 206 -----PHLQANG-----LQNT-WGYNPLAMFAVEPKYAT-- 234  
 OY 585 GTTRIKERTMLOAKKODGMNVIMDVYVNH-----NAAGPTDRTGVLDKIVPM 634  
 DB 235 -NNPLAEKTYVKAHK-AGIEVILDVYFNHSAESEQITPTFSQKIDQT----- 283  
 OY 635 YVORLNE-----TGSVESATCCSDSAPENHFAKLADSLAVWTTDYKIDGFRDLM-- 687  
 DB 284 YWRNDQGRYIMWTGCGMNLNSD-----VGRKVVYDCLRYWVEQCHIDGFRDLATV 337  
 OY 688 -----GYHPKQIILSAMERIKALNPDIYFEGEGWDSNDSFELASQNLKGTGTGTF- 740  
 DB 338 LGRTDPTDNSSQDLTDINKNEPSLQ-NIKLIAEPMD-----IGHGYOYGNNP 384  
 OY 741 -----SPRLRPSVGGCGPFGSDGALRQNOGIGSGAGVLPNELASLSDQVHRLADLTR 794  
 DB 385 SYFAEMNDRFRDLCRFMLKSGE-----IGAFNERFAGSSD----- 421  
 OY 795 GNAGNLIADFWIMDKDAKKGSEIDYNGAPGYADPEEVVYVSKHNDQTLMDMISYKA 854  
 DB 422 -----LEFKKNRL-----PHTTLNFTAHADGFTLLKLDIVSYNQ 453  
 OY 855 SOEADLATFVR-----MQAVSLATVMLGOGIAFDQGGSELRSKFTDS----- 899  
 DB 454 KHNETNGEENRGRNENTSYNHGVBGSESTLS-----EPQKSAVENNRTPAGSGLLMISL 508  
 OY 900 -----YDSGDMFNRYDVS-----LDQNNYVNGMPRISDDGSNVEVTRKEMVATPGE 947  
 DB 509 LANGTPMLLAGDEFGNTQYGNNAAYCQDNE-----ITWLK----- 543

OY 948 AELKONTAFYOELTELKRSPLFTJGDSAVMKRVDFNTG---SDQAGLL-----VMT 999  
 DB 544 -----WANFNEELFTKOT-----IALRKQIGSLKDDQMSDENQVIMLVGEPMT 590  
 OY 1000 VDDGKAKGAS-----LDSRLDGLVAINAPESR 1028  
 DB 591 VEDMONQOTKALQVYLIDNR---WLLITNAKAGQ 621  
 RESULT 10  
 ID WAPA\_BACSU STANDARD; PRT: 2334 AA.  
 AC 007833;  
 DT 01-OCT-1994 (Rel. 30; Created)  
 DT 01-OCT-1994 (Rel. 30; Last sequence update)  
 DT 01-NOV-1995 (Rel. 32; Last annotation update)  
 DE WALL-ASSOCIATED PROTEIN PRECURSOR.  
 GN WAPA OR N17G.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=93302506; PubMed=8316082;  
 RA Foster S.J.;  
 RT "Molecular analysis of three major wall-associated proteins of a  
 RT gene encoding a 258 kDa precursor two-domain ligand-binding  
 RT protein."  
 RL Mol. Microbiol. 8:299-310(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSC1A1;  
 RX MEDLINE=95219088; PubMed=7704263;  
 RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;  
 RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis  
 RT genome containing the hut and wapa loci."  
 RL Microbiology 141:337-343(1995).  
 CC -1- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,  
 CC MOTILITY, SECRETION OR DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED  
 CC INTO THE MEDIUM.  
 CC -1- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE  
 CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE  
 CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED  
 CC MOTIF REPEATED 31 TIMES.  
 CC -1- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME  
 CC SIMILARITY TO THE REPEAT IN E. COLI RHS GROUP OF PROTEINS (RHS-A-D).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L05634; AAA22883.1; -  
 CC DR EMBL: D31856; BAA06656.1; -  
 CC DR EMBL: D29985; BAA06260.1; -  
 CC DR EMBL: D83026; BAA11683.1; -  
 CC DR EMBL: Z59124; CAB15959.1; -  
 CC PIR: S32920; S32920  
 DR Subtilisin, BG10797; wapa.  
 DR Pfam: PF02018; CBD\_6; 1.  
 KW Cell wall; Repeat; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 2334  
 FT DOMAIN 504 869  
 FT REPEAT 504 605  
 FT REPEAT 636 736  
 OR 32 (POTENTIAL).  
 WALL-ASSOCIATED PROTEIN.  
 3 X 101 AA APPROXIMATE TANDEM REPEATS.  
 1-1.  
 1-2.

```

FT REPEAT 769 869 1-3.
FT DOMAIN 1021 2139 31 X 21 AA APPROXIMATE TANDEM REPEATS OF
FT REPEAT 1021 1040 X(4)-G-X(4)-(YF)-X-D-X(2)-G-X(4).
FT REPEAT 1042 1061 2-2.
FT REPEAT 1063 1082 2-3.
FT REPEAT 1083 1102 2-4.
FT REPEAT 1109 1128 2-5.
FT REPEAT 1129 1148 2-6.
FT REPEAT 1150 1169 2-7.
FT REPEAT 1174 1193 2-8.
FT REPEAT 1199 1218 2-9.
FT REPEAT 1219 1238 2-10.
FT REPEAT 1246 1265 2-11.
FT REPEAT 1267 1286 2-12.
FT REPEAT 1290 1309 2-13.
FT REPEAT 1311 1330 2-14.
FT REPEAT 1332 1351 2-15.
FT REPEAT 1373 1392 2-16.
FT REPEAT 1414 1433 2-17.
FT REPEAT 1440 1459 2-18.
FT REPEAT 1481 1500 2-19.
FT REPEAT 1521 1540 2-20.
FT REPEAT 1561 1580 2-21.
FT REPEAT 1598 1617 2-22.
FT REPEAT 1629 1648 2-23.
FT REPEAT 1669 1688 2-24 (APPROXIMATE).
FT REPEAT 1683 1702 2-25.
FT REPEAT 1708 1727 2-26.
FT REPEAT 1732 1751 2-27.
FT REPEAT 1751 1770 2-28.
FT REPEAT 1771 1790 2-29.
FT REPEAT 1793 1812 2-30.
FT REPEAT 1814 1833 2-31.
FT REPEAT 1839 1858 2-31.
FT REPEAT 1859 1878 2-31.
FT REPEAT 1879 1898 2-31.
FT REPEAT 1906 1925 2-31.
FT REPEAT 1927 1946 2-31.
FT REPEAT 1948 1967 2-31.
FT REPEAT 1969 1988 2-31.
FT REPEAT 1983 2002 2-31.
FT REPEAT 2008 2027 2-31.
FT REPEAT 2028 2047 2-31.
FT REPEAT 2051 2070 2-31.
FT REPEAT 2071 2090 2-31.
FT REPEAT 2093 2112 2-31.
FT REPEAT 2120 2139 2-31.
SO SEQUENCE 2334 AA; 258329 MM; B75138CD278BAA3 CRC64;

```

Query Match 318; Score 176; DB 1; Length 2334;

Best Local Similarity 18.98; Pred. No. 0.072; Matches 230; Conservative 163; Mismatches 460; Indels 366; Gaps 55;

```

QY 56 TAVENQA---VHLVDIAGI-----TSSADYSSKNLYLWN---NETCDLSAPV 100
DB 683 TSAEIKAGKXVALHLKDGSGALPINPGTYKKNACCGAKKNYSKILAYKKDGAALSPA 742
QY 101 ADMNDVSTTSGSKYCP-----YVYIPLNKESGCINIVYRDGTDLKIDSLRFAFGDF 154
DB 743 A-----TPALPDIPARPKNVGTLYTNTKSSQGYVNLI-----WEKVONA----- 782
QY 155 TDRIVSYIAGSAVYDSDRADAFRAAGVALAEAHVWDKNTLLMPGGODKPIVRLYSHS 214
DB 783 --KGYKNITNGKETQS-----FVGDND-HWTTQNKINIMPTSEE-----IKAGSY 825
QY 215 KVAADGCKFTDRYLKLTPTTVSQOVSMPRHLSYAFKLPIDNANVDELLOGEFTVAAA 274
DB 826 KLHIDGK-----GRLALDPSVYNNANGNKKCKNYSFTLVAAVDAN-----GETIPTAR 875
QY 275 AEDGILSATQOVAGVLDQAIYAAALSYGAOLADGVTFRYWA-----PTAO-----QV 326
DB 876 -----FNFTEHGCARFIG-----TEEYASTIDIPSGOLNATG 908
QY 327 DVVVVYSAKKVITGSHP-----MTR-----DSAGSAMWOGGS---DLKG 362
DB 909 NVIVNEDDLIDGRGPGIGISRTYNSLDSSDHLFGQGVYADAEISVISTGGAAYIDEDA 968
QY 363 AFYRYAM-----TVYHPOSRRVEQEVTDYPAHSLSTNSESQVVDLNDLSALKP---DCMDN 416
DB 969 TTHHFTKKADQOTQOPRGVYVLELTETADQFL-LKTKDQINAVFNKKGCKLQKVDGHHN 1026
QY 417 LTPHAKTADLAKMTIHESHRLDSAMQOTYPAELRGYLLALTAGDSMNVQHLKTLSA 476
DB 1027 ATV-YTVNDKNQNLATF-----DASGKRLFTYDENGCHVHSITGPKN 1067

```

```

QY 477 SCVTHVELLPVFDLATVNEFSKVDYADIQOPSS-----RLCE---VNSAVNSSEFAGY 525
DB 1068 KKVYV-----SYENDLKKVTPDGTSTYDYSEGRILYKQVSNSTEARPYFTEY 1118
QY 526 CDGSEVEEVLNQLKQSDQDNPOVALNTLVQOTDSYNMGYDOPFHVTEGSAATDPEG 585
DB 1119 QYSGHRLKAIN-----AKKETVYVSYDADKRTL-----LMTOPNG 1154
QY 586 TT---RIKEFRMIAQIKODLGM-----NVIMDVYVNHNTNAGPTDR----- 624
DB 1155 RKQVGYNEAGNPIQVIDADAGLKITNTFKYEGNNVVEDVDPDVCYCGKATSEYQYDKG 1214
QY 625 --TSVLDKIVPWYQ-----RLNETGSV-----ESATCCSDSAPHRM 661
DB 1215 NVTSVKDAYSETEYKRNNDVTKMKDTEGNVTDIAVGDIAVSETDQSSKSSAAVYDK 1274
QY 662 FAKLIADS--LAWTTTDYKIDGFRFDLMGYH-----PKAOLISAMERIKAL 705
DB 1275 YGNOIQSSKDLASATNLKQSGFPAQKSGNNLTASKRRKISYADKSGVLSGSKALEVL 1334
QY 706 NPDIYFEGEGWDSNOSDREFIASQ-INLKGITGTFESDRLRDSYRGGPPDSGALRONQ 764
DB 1335 SQST---SAGFDHGYSS-----ATQVLELPNTYVTLGKIKTDLAKSRAVENIDLBDKQ 1387
QY 765 GIGSGAGVLPNELASLS--DDOVRHLADLRLGMAGLADPFWMI---DKGGAARKG-SEI 818
DB 1388 --KRIOWIHNEYSALAGKNDWTKROITFTTPANAGKAVVYMEVDHDKDKGKAMEDEV 1444
QY 819 DYNGAPGVAADPTEVYVNYVSKHDNOTLMDMISYKASQEADLATRYWQA----- 868
DB 1445 QLEGEVSSSYNPQNSFSFSAIEN---WNVSGASVNSEGFNDSDVSLKAARISASQAGS 1501
QY 869 VSLATVNLGOG-----IAPDOGSSELRKSFTRDSYSDGDMENRVDYSLQDN----- 916
DB 1502 VTKQTVVLGOSANDKPYVLTILTGMSKASVYFTDEK-----DYSIQAVNTYADGS 1551
QY 917 --NYNVGMPRIISDGSNVEYITRKKNVATPGAEELKQM---TAPQELTELKSSPLF 970
DB 1552 TGIYNAKPSGTQEMNAAVYIPKTRPKINVDISILEQKATGATWDDRLTEGS--LL 1609
QY 971 TL---GDGSAVMKR-----VDFRNTGS-----DDOAGLLVWTVD 1001
DB 1610 TKSTYDNGNVTYKEDELGYATSTYDCTCKTSEFDAGKERTYTYDDADQILNTMFLS 1669
QY 1002 DG-----MKAGASLDRLDGLVVAINAAPESRTLNEFAGETQLQSAIQ 1044
DB 1670 NGTISILSHYDKEGNEVSKITRAGADQTYKFEYDMGRLVTTTDPGLGNVLAISEYDANSLT 1729
QY 1045 QTAGENSLANGVOIAADGT 1063
DB 1730 KTISP--GNEVSLSDGT 1746

```

RESULT 11  
 YP62\_MYCTU STANDARD; PRT; 580 AA.  
 ID YP62\_MYCTU  
 AC Q10769;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOHETICAL 64.1 KDA PROTEIN RV1562C.  
 GN RV1562C OR MTCY48.03.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID-1773;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE-98295987; PubMed-9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,



```

Db 77 TIVPENLEGLYKY--EITPDGRVLK---ADPAFYSELRRHTASIV-----YDLKGY 127
QY 413 GMDLTPHNAOKTADLAK-WTHIESITRLSANDQVPELRRKYALATAGSDNMVQHL 471
Db 128 EWNDSFMRKRRRRRIYDOPVYIELH---FGSMK---KPDREFYTYREMADELIPY- 179
QY 472 KTLSASGVTHVELLPVDFLATVNEFSKVDADIOOPFRLCEVNSAVSSSEFAGCDSGST 531
Db 180 --VLERFTHIELPL-----VEHPLDR-----200
QY 532 VEEVLNOLKSDSQDNPOVALNTLVAQDTSYNGMDPFHYTYPEGSAYADPEGTTRIKE 591
Db 201 -----SMGYGTGYSTJSRYGTP-----HD 221
QY 592 FRTMIOAIKODLGMNVMIDVYVNH--TNAAGPTDRTSVLDKIYVWYQRLNETTGSYESA 649
Db 222 FMYFVDRCHQ-AGIGVYIDWPVGFHCDAHG---LYMFCDAPTYEVANEXKRENYWGT 276
QY 650 TCCSDSAPRHRMFALIASLAVWTTDYKIDGFRFDL---MGYHPKAOIL-----SAMESI 702
Db 277 ANFDLGRPEVRSF--LISML-FPLEYHYHDGFVDAVANMLYPPNDRLEYENPYAVEFL 333
QY 703 KALN-----PDIFYEGE--GMDSNQDREFEIASQINLKGITGFSRDL-----744
Db 334 ROLNEAVFAYDPVNMIAEDSTDPRTVAPTVDGLGPNK-NMGMMDMLKXMETPPH 392
QY 745 -RDSVRGCGPDSGDALRQNDGIGSGAVLPNELASISDDOYRLADLTRLGAGNLAADF 803
Db 393 ERKYNHNSVSSLLYASEN-----FILP-----FSHBEVVI-----424
QY 804 VMIDKCAAKKSEIDYNGAPGYA--ADPTEVNVVSKHNDNTLMDMISYKASQEADL 860
Db 425 -----GKSLINKRPGSYEEKFAOLRLIXGYMAHPKGLPM-----462
QY 861 ATRVMOAVSLATVLMOGIAFDQGSSELLRSKSTRTSDSGDFWFRVYSLQDNVYV 920
Db 463 -----GSEFAO-----FDEMFAELDMVLFD-----484
QY 921 GMPRISSDGSNVEYITRVKENVATPGEAEIKOMTAFYOELTELKSSPLTLGSGSAVMK 980
Db 485 -----FELHRMDEYV-----KOLACY-----KRYPFELDDHPRGFE 519
QY 981 RVDFRNTGSDOAGLWVTVDDCKAKASLDSRLDGLVVAIN-----AAP 1025
Db 520 WIDVHNA---EQS---IFSIFRKGKES-----DVLVIYCNFTNOAYDXYKVSPLIAP 567
QY 1026 ESRTLNFAGETLQLSAIQGTAGENSILANGVQIADG-----TVTLPAWSAVL 1074
Db 568 YREVLNSDAE-----FCGSGHVNGKRLPAPSEPFHGKPYHYVMTLPPGISIL 616

RESULT 13
GLGB_MYCTU STANDARD; PRT; 731 AA.
AC 010625;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BRANCHING 1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18) (GLYCOGEN
DE BRANCHING ENZYME).
GN GLGB OR RV1326C OR MTCY130.11C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN SEQUENCE FROM N.A.
RP STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

```

```

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.C.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC GLYCOGEN.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESES.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC DR EMBL: 273902; CAA98090.1; -.
CC DR Tuberculist: RV1326C; -.
CC DR InterPro: IPR000461; -.
CC DR Pfam: PF00128; alpha-amylase; 1.
CC KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
CC FT ACT_SITE 411 411 BY SIMILARITY.
CC FT ACT_SITE 464 464 BY SIMILARITY.
CC FT ACT_SITE 532 532 BY SIMILARITY.
CC FT SQUENCE 731 AA; 81729 MW; EE2BREF65352617 CRC64;

Query Match 3.0%; Score 166.5; DB 1; Length 731;
Best Local Similarity 21.3%; Pred. No. 0.043;
Matches 163; Conservative 96; Mismatches 273; Indels 233; Gaps 41;

QY 202 DKPIVRLYTHSSKVA--DGBGKTRDKLTPTVSQOVSMPRHLISYAAFLPDNAN 260
Db 41 DHTYIRAFRPAVEVALVKGDRFSLOHLD-----SGLEFAVALPVDL 83
QY 261 VDELLGGEYVAIAAEGLISATQ--VOTAGVLD-DAVAEAAALSY---GAO-----LA 310
Db 84 IDYRLQ---VTYEECEHRYVADATVRLPTLAGEVULHLPAGEHRIEVLICAPRSTTTA 140
QY 311 DG---GVTFRWMAPTAOQDVV-----YVSADKRYIGSHPTRDSASGAMSQGSDDL 360
Db 141 DGVVSGVSPVAMVAMNAKGVSLIGEFNGMNGHEAPMRLG-----PSGVWELFWPDPF 192
QY 361 KGAFYRYAATVYVHPOSKRVEQYETDPRVHSLSTNSEYSOVVDLNSALKPRIDMDNLTMP 420
Db 193 CDGLYKFRV---HGADGVV--DRADPFAEGTEVPQTAGRVVTSDDYTWDDDW---MAG 244
QY 421 HAOKTKADLAKMTIHESHITDLSAMDQTPAELGKAYLALTAGOSNNVHOLKTLISAGVT 480
Db 245 RALRNPVNEA-MSTYEVH---LGSW-----RPGLSYROLAHELDYIDQ-----GFT 288
QY 481 HVELLPVFDLATVNEFSKVDADIOOPFRLCEVNSAVKSEFAGYCDGSGTVEVLNQLK 540
Db 289 HVELLPVAE-----HPFAC-----302
QY 541 QSDSQDNPOVALNTLVAQDTSYNGMDPFHYVPEGSYATDPEGTTRIKEPRTMIOAIK 600
Db 303 -----SMGYQVTSYAPTSRFGT-ID-----DFRALVALH 332
QY 601 QDLGMNVMIDVYVNH--TNAAGPTDRTSVLDKIYVWYQRLNETTGSVESATCCSDSA 656
Db 333 Q-AGIGVYIDWPVGFHCDAHG---LYMFCDAPTYEVANEXKRENYWGT-----GR 385
QY 657 PEHRMFALIASLAVWTTDYKIDGFRFDL---MGYHPKAOILSAMESIRKALNPDIYFEG 713
Db 386 PEVRNMF--LVANML-VWLOGFHIDGLRVDVAVSMILYIDYSRPBGW-----TPNVH---433

```

[illegible]

14	RESULT		
1	GLGB_HAFTN		
ID	GLGB_HAFTN	STANDARD:	PRT: 730 AA.
AC	P45177;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-NOV-1995 (Rel. 32, Last annotation update)		
DE	1.4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18) (GLYCOGEN BRANCHING ENZYME).		
DE	GLGB OR H11357.		
OS	Haemophilus influenzae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Haemophilus.		
OX	NCBI_Taxid=727;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=RD / KW20 / ATCC 51907;		
RX	MEDLINE=95350630; PubMed=7542800;		
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirtness E.F.,		
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,		
RA	McMenemy K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,		
RA	Scott J.P., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,		
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,		
RA	Wettersack T.R., Hanna M.C., Nguyen D.T., Staudk D.M., Brandon R.C.,		
RA	Elze L.D., Fitchman J.L., Fuhrman J.L., Geoghagen N.S.M.,		
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,		
RA	Venter J.C.;		
RT	"Whole-genome random sequencing and assembly of Haemophilus		
RT	influenzae Rd.";		
RL	Science 269:496-512(1995).		
CC	-1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF		
CC	GLYCOGEN.		
CC	-1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.		
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO		
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL: U32815; AAC23004.1; -.		
DR	TIGR: H11357; -.		
DR	InterPro: IPR000461; -.		
KW	Pfam: PF00128; alpha-amylase; 1.		
FT	Glycogen biosynthesis; Transferase; Glycosyltransferase.		
FT	ACT_SITE 405 405 BY SIMILARITY.		
FT	ACT_SITE 458 458 BY SIMILARITY.		
FT	ACT_SITE 526 526 BY SIMILARITY.		
SO	SEQUENCE 730 AA; 83820 MW; 58957531/F53769A CRC64;		

Query Match	2.9%	Score 163	DB 1	length 730
Best Local Similarity	19.2%	Pred. No.	0 068	
Matches 153	Conservative	86	Mismatches 244	Indels 312
				Gaps 39

QY	303	LS7GADL	-----ADG	--GVTEVMACTAQDUDV	---VYSARKVIGSHPRK	343	
Dd	113	LSGSMLRPEYVLCGAFHFMEDCGVSGVNFRLMAPNARVSIYGDFFVMDGR <th>----</th> <th>HPMR</th> <td>168</td>	----	HPMR	168		
QY	346	DSASGAMSMGCGDGLGAFRYMYTYHHQSKRVEYETVDYHSLNSTSEYQVDLN	405				
Dd	169	HSSGSGWELFLPRASLDQYKFEELDCHGNLR <th>----</th> <th>KADFAFS</th> <td>209</td>	----	KADFAFS	209		
QY	406	DSALRPGDMNLT <th>-----MPHAQKTKADLAK</th> <th>--MTIHESIIRLSAWDQVPAELRGK</th> <td>456</td>	-----MPHAQKTKADLAK	--MTIHESIIRLSAWDQVPAELRGK	456		
Dd	210	SQURPPTASQVVALPRVYEMTEARK <th>---KANQGNQPSIYEYH</th> <th>---LGSRRULENNFMD</th> <td>264</td>	---KANQGNQPSIYEYH	---LGSRRULENNFMD	264		
QY	457	YLALTAGDSMMVOHLKTLASGVTHVELLPVEDLATVNEFSKRVADIDQPSRLCEVNSA	516				
Dd	265	YDQJLA <th>---DELIFVYKEM</th> <th>---GPTHIEFLP</th> <th>---LSEF</th> <td>294</td>	---DELIFVYKEM	---GPTHIEFLP	---LSEF	294	
QY	517	VKSESEFAGYCDSDSTBEVNLNQKQSDQDNPNQVALNLTVAQTSYNGXDPFHYTYPE	576				
Dd	295	-----DGS	-----WGTOPCLLYSPT	309			
QY	577	GSVATDECTTRIKERFTMIOALKODIGNVIMDVYNN <th>-----TNAAGPTDTSVLDR</th> <td>IV</td> <td>632</td>	-----TNAAGPTDTSVLDR	IV	632		
Dd	310	SRFS <th>---PEA</th> <th>---PRRLTKR</th> <th>---AHGAGIVLIDWV</th> <th>RGHFRPDTGIVAFDGTALYHED</th> <td>361</td>	---PEA	---PRRLTKR	---AHGAGIVLIDWV	RGHFRPDTGIVAFDGTALYHED	361
QY	633	P <th>---WYQRLNE</th> <th>---TTGSVESATCCSDSAPENHRMAKLIADSLAVWTTDYKIDGFRFDLM</th> <td>687</td>	---WYQRLNE	---TTGSVESATCCSDSAPENHRMAKLIADSLAVWTTDYKIDGFRFDLM	687		
Dd	362	PREGHODMWTLLYNGRNEVKNFLSSNA <th>-----LYWLERFVGDGIRVD</th> <td>405</td> <td></td>	-----LYWLERFVGDGIRVD	405			
QY	688	GIYHKAQLSAMERIKALANDIYFPEBGWDSNDSDFELASQINLKGTGIGFSRLRDS	747				
Dd	406	-----	-----AASMI	-----YRDY	415		
QY	748	VRGGPDPDSDGLARONOGIGSAGVLPNELASLDDOVYHNLADTLRIGMAGMIADPFVMD	807				
Dd	416	SRAGEN <th>-----IPNQGGRENLEALIEPLKHTMWKIHSEMAAISIA</th> <td>457</td> <td></td>	-----IPNQGGRENLEALIEPLKHTMWKIHSEMAAISIA	457			
QY	808	KDCAKKG <th>---SEIDYNGAPG</th> <th>-----GYADPLEYVN</th> <th>---YVSKHNOTLMDMITYAS</th> <td>855</td>	---SEIDYNGAPG	-----GYADPLEYVN	---YVSKHNOTLMDMITYAS	855	
Dd	458	EESTSFAGVYHPSHNGSIGCFNFKMNGMWNMDTLAYMKLPIROYUHNHMTFGAM <th>---YOYS</th> <td>516</td> <td></td>	---YOYS	516			
QY	856	QEADLA <th>---TRVRQAVSLTYVLGQ</th> <th>-----GIARDQGSSELLRKSFTPROSY</th> <td>900</td>	---TRVRQAVSLTYVLGQ	-----GIARDQGSSELLRKSFTPROSY	900		
Dd	517	ENFPLPSHDEYVNGKSLCKMRPGDTWQFANLRAYIGMYKPEKKLL <th>-----BNGNEF</th> <td>572</td> <td></td>	-----BNGNEF	572			
QY	901	DSCDMFN <th>---RVDYSLQDNMYNYGMPRISDDGSNEYITRVKEMVATPEGL</th> <th>---KONTAF</th> <td>956</td>	---RVDYSLQDNMYNYGMPRISDDGSNEYITRVKEMVATPEGL	---KONTAF	956		
Dd	573	AQGEWYVESLDMLFDENIGGWMK <th>-----GVLLAYVDLNOI</th> <td>611</td> <td></td>	-----GVLLAYVDLNOI	611			
QY	957	YQELTELKRSPLYTLGDSAVMKRVDFRNTGSDQOAGLLVMTYDDMKFAGASLDSRL <th>---</th> <td>1014</td> <td></td>	---	1014			
Dd	612	YQ <th>---KNRPTEL</th> <th>-----DNSPEGDMVLVDDAANSVLAFFERRSN</th> <td>649</td>	---KNRPTEL	-----DNSPEGDMVLVDDAANSVLAFFERRSN	649		
QY	1015	DGLVVAINAPESR	1028				
Dd	650	GERITIVSNETPVPR	664				
RESULT 15							
APU_THETU							
AC	P38536	STANDARD	PRT	1861	AA.		
DT	01-OCT-1994	(Rel. 30, Created)					
DT	01-NOV-1995	(Rel. 32, last sequence update)					
DT	15-JUL-1999	(Rel. 38, last annotation update)					
DE	AMYLPOULIUNANASE	PRECURSOR (ALPHA-AMYLASE/PULIUNANASE)	(PULIUNANASE				
DE	TYPE II)	(INCLUDES: ALPHA-AMYLASE (EC 3.2.1.1) (1.4-ALPHA-D-GLUCAN					
DE	GLUCANOHIDROLASE)	PULIUNANASE (EC 3.2.1.41) (1.4-ALPHA-D-GLUCAN					
DE	GLUCANOHIDROLASE)	(ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)]					
AN	AMB.						

OS Thermomicrobacter thermosulfurogenes (Clostridium  
OC thermosulfurogenes).  
OC Bacteria: Firmicutes; Bacillus/Clostridium group;  
OC Thermomicrobacter group; Thermomicrobacterium.  
OX NCBI\_TaxID=33950;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-DSM 3896 / EM1;  
RX MEDLINE-94252998; PubMed-8195085;  
RA Matuschek M., Burchard G., Salm K., Bahl H.;  
RT Pullulanase of Thermomicrobacterium thermosulfurogenes EM1  
RT (Clostridium thermosulfurogenes): molecular analysis of the gene,  
RT composite structure of the enzyme, and a common model for its  
RT attachment to the cell surface.  
RL Bacteriol. 176:3295-3302(1994).  
CC -1 CATALYTIC ACTIVITY: ENDOLYOLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
CC -1 CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDOLYZES  
CC (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO  
CC FORM MALTOSE.  
CC -1 SUBCELLULAR LOCATION: CELL-BOUND. IT C-TERMINUS MAY SERVE AS AN  
CC S-LAYER ANCHOR.  
CC -1 PTM: GLYCOSYLATED.  
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC -1 SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -1 SIMILARITY: CONTAINS 3 S-LAYER HOMOLOG (SLH) DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL: M57692; AAB00841.1; -  
CC DR InterPro: IPR000461; -  
CC DR InterPro: IPR001119; -  
CC DR InterPro: IPR001777; -  
CC DR Pfam: PF00395; SLH: 3.  
CC DR Pfam: PF00128; alpha-amyase: 1.  
CC DR Pfam: PF00041; fn3: 2.  
CC DR PROSITE: PS01072; SLH DOMAIN: 3.  
CC KM Hydroxylase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;  
CC KM Multifunctional enzyme; Glycoprotein.  
CC FT SIGNAL 1 35  
CC FT CHAIN 36 1861 AMYLOPULLULANASE.  
FT ACT\_SITE 628 628 BY SIMILARITY.  
FT ACT\_SITE 657 657 BY SIMILARITY.  
FT ACT\_SITE 734 734 BY SIMILARITY.  
FT DOMAIN 928 1018 FIBRONECTIN TYPE-III.  
FT DOMAIN 1157 1248 FIBRONECTIN TYPE-III.  
FT DOMAIN 1681 1739 SLH 1.  
FT DOMAIN 1740 1803 SLH 2.  
FT DOMAIN 1804 1861 SLH 3.  
FT COMPLET 1734 1734 D -> E (TN AAB00841).  
SQ SEQUENCE 1861 AA; 206104 MM; 06C23070E453B574 CRC64;  
  
Query Match 2.8%; Score 160.5; DB 1; Length 1861;  
Best local Similarity 19.3%; Pred. No. 0.38;  
Matches 216; Conservative 148; Mismatches 402; Indels 351; Gaps 56;  
  
OY 3 RYRNALVGLSLVLLSGDNGSSSSSSSGN--BDTPNDQDVVRLPDAVAGEAVAVEN 60  
DB 882 KFLRDGVTFKDLI-----NNNVYSISNCOIVIDVPAWSCVM-----ISDDGDLTAPQA 932  
OY 61 QAVIHL-----VDIAGITSSAADYSSKNLYLMNNE--TCDLASAPVAD--WNDVSTP 110  
DB 933 PSNVAVTSGNGKVDLSMLSDGATGY--NIYSSVEGGLYERIASNVETTFEDANVTN 989  
OY .1LJ SGGDKGPTWVIVPLNKESGCIIVTR-----DGTDXLIDSDLRVA 150

DB 990 GLKRYVAISAIDELGNESG:ISDAVAVPAVPGVGLTVGSDNHIIIGVDPEDYAEV 1049  
OY 151 FGD-----FTDRIVSVING-----NSAVYDSRADAFRAAGVALAEAWDCKTLLP 200  
DB 1050 WADGILFNTSGCGPRNMLAQOLQKYYSGTVYDS-----VYGSVYNSVYGVDSGFTVYNA 1102  
OY 201 QDKPIYRLYYSHSSKVAADGEGKFTDRY-LKLPTTVSO-QVSNRPPLHSSYAAFKLPDN 258  
DB 1103 Q-----YV-----GDIGNNDQYKASFPTDKIGQMEYLMRF-----SDN 1135  
OY 259 ANVDLLQGTVAIAAEDGILISATQVOTAGVLDVAIAAALSLSGADLADOGTVFRV 318  
DB 1136 QGQDWITTTSTLSTFYVPSDDL-----KTPAPYLNOPTGSSSR-----VSLT 1177  
OY 319 WAPTAQVDVAVVYSAADKVGSHPMTRDSASGAMSGSDLGAFFRYAMTVYHPQSRK 378  
DB 1178 WNPSTQVNGVYDYEIRHSDGCTFNKIATVENEYNTIDTSVINGVYTVYKVV----- 1229  
OY 379 VQYEVTPYAHSLSTNSESQVVDLNDLAKPD----- 412  
DB 1230 -----AVDLSEFRTESNV-----TIKPDVPIKVIYFVTVPDYTPDAVNLATGF 1274  
OY 413 -----GMDLTPHAKK--TKAD-----LAKMTIHE-----SHIRLSAMDQVPAELRKY 458  
DB 1275 PNATWD-----PSAQOMTKIDNNYTSITLTLDGCTOLEYKYAR--GSWDKYEKDEYGEFA 1328  
OY 459 A-----LTAGDSNMVQHLKTLASGVTHELLPFEDLATVNEFSDVADIQPFSLCE 512  
DB 1329 SNKKVTVIVNGGNM-----TIN-----DVIYMRDIPRIYSPSSNM-T 1367  
OY 513 VNSAVSSSEFAGYCDSCSTVEEVLNLOKSDSDNPQVALNTLVAOTSDYNNGYD--PF 570  
DB 1368 VDSNISITMEKYGKTYKAKV--TIN-----GDS-----FVQDKNGVFTDVLNGLVKNIKI 1417  
OY 571 HYVPEGSYATDDECTRIEFTPMQAIKODLGMVIMVYVNHNTAACTPDRTSVLDK 630  
DB 1418 HVEPNDGSYGNDOG--RIIE-----LTKDEIVYHOENNSGSGCTGNN----- 1460  
OY 631 IVPWYQRLNETTGSVESAT-----CCSDSAPHEHMEFKLILADSLAVTTDYKIDGFRFDL 686  
DB 1461 -----TSTSGSNSSSTGSGSTGTSITSNSNTSNTIGVITKKGNVITLTLDA 1511  
OY 687 MGYHPRKQILSANDRIKALNPDIYFREGWDSNOSDFELASQINLKGITIGFSRLRD 746  
DB 1512 -----GKAKDLIVSKDKKVVFDITITIGEGQKVVQISKDIIDTSAAANGKDIVKSDNASI 1567  
OY 747 SVRGCGPFDGDLRONQIGSGAGVLPNELASLSD-----QVRHLADLTTRLGMA 797  
DB 1568 AL-----TKDALNQNG--LQNGVY-----SIKNGKPNVNTVYSLSNVVDITISGIS 1613  
OY 798 GNLADPVMIDKGAANKGSEIDYNGAGGYAADPTEY-VNYVSKHNDQTLMDISYKASQ 856  
DB 1614 GNV-----TLAKPVEVTLNISK--ANDPRKVAAYVYVNPNTNQ--MEVYGGKVDA 1658  
OY 857 EADL-----ATRVKQAV-----SLATVMLGQGIAPFOQSESLRSKSF 896  
DB 1659 SSGTITFNATHFSQYAAFEYDKTFNDIKDMANQVLEVLASRIHVEGMDTQYEPNKTVT 1718  
OY 897 RDSY-----DSGDMF--NRVDYSIQDNNYVNGMPRIDDSQSN-- 931  
DB 1719 RAFTFAMILRLINKDFTYSGERSDYKSGDMYANALIEA-----YKAGI--ISDGKNAR 1771  
OY 932 -YEYITRVKEMVATPGAEALKQMTAFYQOELTELKSS 967  
DB 1772 PNDSTIR-EEWTAIAMRA-----YEMLTQYKEEN 1799

Search completed: April 22, 2001, 09:12:11  
Job time: 370 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 09:10:53 ; Search time 74.41 Seconds  
(without alignments)  
1718.503 Million cell updates/sec

Title: US-09-262-126c-4  
Perfect score: 5633  
Sequence: 1 MURYSNALVLGSLVLLSGC.....AVLELPQGAAGALFVSSK 1091

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_15:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	1667.5	29.6	1798	2	Q9KZ11	
2	1503	26.7	964	10	Q41386	
3	1432.5	25.4	10	081638	081638 spinacia ol	
4	1424.5	25.3	904	10	048541	048541 hordeum vul
5	1420	25.2	910	2	Q9KX58	Q9KX58 hordeum vul
6	1405.5	25.0	910	2	Q9KX50	Q9KX50 delinococcus
7	1386.5	24.6	988	10	064454	064454 oryza sativ
8	1326.5	23.5	986	10	P93416	P93416 oryza sativ
9	722	12.8	668	2	P71095	P71095 ferdidobact
10	692.5	12.3	849	2	Q9KX85	Q9KX85 ferdidobact
11	692.5	12.3	1938	2	P70983	P70983 bacillus sp
12	685	12.2	825	2	Q59319	Q59319 caldocellum
13	647.5	11.2	718	2	Q34387	Q34387 bacillus su
14	630	11.2	718	2	Q69008	Q69008 thermus sp.
15	605.5	10.7	1072	2	Q9K6N1	Q9K6N1 bacillus ha
16	593.5	10.5	717	2	Q9K7U5	Q9K7U5 bacillus ha
17	367.5	6.5	715	2	Q9KY03	Q9KY03 streptomyc
18	362	6.4	715	2	Q9X947	Q9X947 streptomyc
19	355	6.3	782	2	Q9LIE9	Q9LIE9 streptomyc

ID	Q9KZ11	PRELIMINARY:	PRT: 1798 AA.	
AC	Q9KZ11			
DT	01-OCT-2000 (TREMELREL. 15, Created)			
DT	01-OCT-2000 (TREMELREL. 15, Last sequence update)			
DT	01-OCT-2000 (TREMELREL. 15, Last annotation update)			
DE	PURATIVE BI-FUNCTIONAL PROTEIN (SECRETED ALPHA-AMYLASE/DEXTRINASE).			
GN	SCIOB7.21C.			
OS	Streptomyces coelicolor.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=1902;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-A3(2);			
RA	Seeger K.U., Harris D.;			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-A3(2);			
RA	Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-A3(2);			
RA	Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,			
RT	Kinashi H., Hopwood D.A.;			
RT	A set of ordered cosmids and a detailed genetic and physical map for			
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.;			
DR	Mol. Microbiol. 21:77-96(1996).			
DR	EMBL: AL355752; CAB90874.1;			
SO	SEQUENCE 1798 AA; 194106 MW; 43D13868FAC56CA CRC64;			

Query Match 29.6%; Score 1667.5; DB 2; Length 1798;  
Best Local Similarity 36.4%; Pred. No. 4.4e-97;  
Matches 397; Conservative 166; Mismatches 391; Indels 137; Gaps 28;

QY 34 DPPNDQVYV-RLPDVAVGEA-VYAVENQVYHVDINGISSSA-----ADYSSK 83  
DB 793 DDXGNDVADRTIDVTETGEVWIDGEOVLTERPEYPAODTTKAVLHYKRAADGNTDGM 852

Qy	84	NYLNNMETODALASAPADNDVSTPPSSDSDXYGFWYPLRLKEGGCJNVLYRGGTDLI	143
Db	853	GLHWGCD-----AANPTDMAK-PIQRPVFTDYGAVFVFRPLIDGASSLSLYVHNGDEKDL	905
Qy	144	DSDLRV-----AFGDFTRDYVSYAGNSANYDSRADAFAAFGVALAEAHWDKN	193
Db	906	PTDQALDLKADGHEWLLSGDEDHLLPQAPGSAAVD-----LTTSSKAWMDRN	954
Qy	194	TLHPGQODKPIRLLYSHSKAAADGECKFT---DRYJLKLPTMVSQOVSRRPHLSY	250
Db	955	TYVANGSDGAASVOLLAIRSGSTTAEG-GRLTGDDQOQIRLTKSLTMDQAKAPRHLKES	1013
Qy	251	AAEFLP--DNANVELLOGETVAIAAEEGJLISATOVQTAGVLDDAAYAEAAELSYGAQ	308
Db	1014	TAMSVDRDRDRREALRGQVVAHQRTATGALLATGVTAGALDLADYADRATRADGPV	1073
Qy	309	LAOGGVFRYMAATTAQOVNVVYSAADKXVYIGSHIPTRDSASAMWOGSDLKAFYTXA	368
Db	1074	FRIGRPLTSLWAPPAO--DKLEIGIRTV---RMRKRDATGWSVTGPKSMKGAARYA	1127
Qy	369	MTVYHPQSRVEQEVYTDPAHSLSTNSEYSOVVDLNDLSALRPGDQMDLTPMHAQTKAD	428
Db	1128	VKWAAPAGEVYVYNTKYTDPRSLATLTDSERSLVDLDDGSLAPRMEHYRRPEA---VP	1183
Qy	429	LAKKTTHEIIIROLSAMDOVPAELRGKTLALTAGDSNNVOHLKTLASGVTHWELPVE	488
Db	1184	LQDAQIOELRHOFVSADRDYADADHNGTGLAFIDKKSDGSRHLRELAAGTSYVHLPAF	1243
Qy	489	DLAVNFEFSKVDADIOQPFRLCEVNSAKSSEFGYCC---SGSTVEEVLNQLKSD	543
Db	1244	DIAITPE-----RKSEQITPDCDLALPAGSE-----KO--	1272
Qy	544	SQDNPOVALNTLVAQDTSYNMGYDFPHYVTPEGSYAIDPEGTTRIKBFRTMIAIKODL	603
Db	1273	-----QECVTAKAAKDAFNMGYDYHYHTVEEGSYATPDGTEFTVEFRKKVYALNED-	1324
Qy	604	GMNYIMDVVYNNHNAAGPIDRTSVLDKIYPMWYORLNETGTSVEBATSQCSOAPREHFA	663
Db	1325	GLRWYMDVYNNHNPASQA-KTISVLDRIYPGYORL-LADGSVANSTCCSNATFENAMG	1382
Qy	664	KLINDLAVNTTYTKIDGFEFDMJGHPYAOQLISAMERIKAL-----NPDITFEFEG	715
Db	1383	KLAVDSVYVTAKEKXKVDGFRFDMJGHPANILAVAKALDELTLRKGDVCKKIYLYGEG	1442
Qy	716	WDSNQ---SDREFASQINLKGTCIGTFESDRILDSYRGCGPSPDSDALRQNGICGSCGV	772
Db	1443	WNFEEDVADDAEFVQATOKNNAQGIATFEDBRADYARGGGPEDEBPV---QGFASGYLT	1499
Qy	773	LPNELASLSD-----QVRLADLTRLGMAGNLADPVMIDKGAKKSESIDYNCAPGY	827
Db	1500	DPNNSYANGTEADQKALHYQDLKVGCLGNLADYSPFDTHCKKEKSEVDYNCAPGY	1559
Qy	828	AADETEVYVYSKHNDOTLMDLSTYKASEADLATFRVQAVSLATVMLGOGIAEDQGS	887
Db	1560	AAAGADLAAADAHNDNESELDALTEFLPKPQTSADDBARQVYLAAMATATLSQPALSGT	1619
Qy	888	ELLKSKFETSDSDSGMFRKRVYSIOD--NNVNYGMPRLSDSGSNVEYITRKENVAPRG	946
Db	1620	DLLSKSLDRKSTSDSGMFALHMCADGNGGFRGLPRADNDQWPKAPKPLGAVEY-G	1678
Qy	947	EAEKQMTAFYOEITELDKSSPLETIGDGSAAVWRVDFNNTSSDOAGLLVTVYDDGPKA	1006
Db	1679	CPQITGASAAVYRDLRLRSGEQFSLATAAAROSRLSPPLSGEDRTPGVITM-----	1730
Qy	1007	GASIDSLRDLGLVAAINAAP--SRTLNEFAGETLQLSAIQOTAGENSILANGVQIAADGV	1064
Db	1731	-----RLGDLVYVFNATPERQORVGAAAGCGYRLHPVQAAGADAVVKEASATAKTGT	1784
Qy	1065	TLPAWSVAYLE 1075	
Db	1785	TVPARIVAFQ 1795	

RESULT	2			
ID	041386	PRELIMINARY;	PRT;	964 AA.
AC	041386			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-MAY-2000	(TREMBLrel. 13, Last annotation update)		
DE	PULULANASE.			
CN	PULSPO.			
OS	Splachna oleracea (Splachn).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Euphorbiales; core eudicots; Caryophyllidae;			
OC	Caryophyllales; Chenopodiaceae; Splachna.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=8 WEEKS OLD LEAVES;			
RA	Renz A., Schmid R., Kossmann J., Beck E.;			
RL	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; X83969; CA58803.1; -			
DR	MENDEL; 16403; Sp1ol;2480;16403.			
DR	INTERPRO; IPR000461; -			
DR	PFAM; PF00128; alpha-amylase; 1.			
FT	CHAIN 65 964			
FT	SEQUENCE 964 AA; 106551 MW; 7CONSCEOD4ALEFSB CRC64;			

Query Match	26.7%	Score 1503	DB 10	Length 964
Best Local Similarity	36.2%	Pred. No. 4,5e-87		
Matches	365	Conservative 169	Mismatches 324	Indels 150
			Gaps	33
QY 123	PLNKESGGINVIVRDGDTKLDIDSLRVAFGDFDTRTVSIAGNSAVYDSRADAFRAAGCV	182		
DB 53	PLKKSSSC-----FCCSMAVEVGSASVSOSBLS-----L	84		
QY 183	ALAEHWVDKNTLLM-----PGSQDKPIRYRLXYSHSXYAAADSEGFETR-----YTKL	231		
DB 85	NSCRATWPSKSTTFANNVNDIGNS-----YLFASKTAA--LKFTDGLGEGYVVKIKL	134		
QY 232	-----PFTTVSOQVSMRFPHLDSYAAREKLPDNAVVDLQETVAIIMAAEDGILISATO	285		
DB 135	DKDGGLPANTYTE-----KEPHIRGYSAFKAAPATLDVDSLKQGLVAAPASADGACRNMTG	190		
QY 286	VQTAGVLDATAEAAEALSY-----CAQLADGCVTFRVWAPTAQVDVYVYSADKKVIGSH	341		
DB 191	LQLPVEDELY-----SYDPLGDAVESENITSLYLAAPTAQVVASISF--KDPSCGE	240		
QY 342	PMTR---DSASGAWSMOGSDLKGAFFYRYAMTVYIPQSRKVEQYEVTDPYAHSLSTNSEY	398		
DB 241	PLQVQVLIENSGNVSAVAGVRMGEGCYUYEIVYHNHSTLRIEKSFAIDPYANGISADVYKR	300		
QY 399	SOYVULNDSALKPDGWDNLT--MPPAOKTKADLAKMTTHESHTIROLSAMDOVTVAELRGK	456		
DB 301	TLLADLSEETLKEGEMENLADKPHILSP---SOLSEYELHRIREFSAVDLTVHVDLRLRG	356		
QY 457	YLATPAGCSNMVQHUKTISASGVTVELLPYEDLTATVNEFSPKAVADIQOPFRLCEVNSA	516		
DB 357	YLAFTSQBSAGCNHLEKLSAAGLTIVHLLHSPQFAEVD--DK-----	397		
QY 517	VKSSEFAGYCDSGSTVEEVLNQLKOSDSODNFOVALMTLVAQTSYMGKIDPFHYTPE	576		
DB 398	-KKMF-----VDTRKFEFTLPPDSEEOOQITAIIDEGYMGNGVNPVLMGTPK	444		
QY 577	GSYATPDGGTTRIKREFRTMIOIKODLGMNVIMDVYVYHNTNAGCTDRSVLDKIVPMY	636		
DB 445	GSYATPDGPPCTIIEFRKMYQALNR-IGIRAVYLDVYVYHNLNSGSDNSVLDKIVPGY	503		
QY 637	QRLNFTTSSVESATCCSDAPAEHRFAFKLADSLAVWTTDYVIDGFRFRDLYGHRKAOL	696		
DB 504	LR-RNDGAIENSTGVNDYASHEFVVERLIDDLKHNMAVNYVDGFRFRDLMGHIMKHMYV	562		
QY 697	SAMEIKIALNDP-----IYFBEGBDMSNO---SDREFLISQNLKGTGIGTISDRLR	745		

```

Db 563 KATNMLQGLSKNIDGVESSSTYLIGEEMDFGEVANNAKVANSQNLGEGTIGSFENDRIR 622
QY 746 DSVRGCGPDSGDALRQNOGIGSGAGVLPNE--LASTSDOVRHLA-DLTRLGAGANLA 801
Db 623 DAVLGCGPFE--GPRPQ--QGYVTGLSLQPNHDHSGKANADMLAVAKDHQIVGAGNLR 678
QY 802 DEVMIDKCAAKKSGEI-DYNGAPGCIYADPTTEVNVYVSKHDNQTLMDSIKSASOEDL 860
Db 679 DYILNCGCKOVKGSSEVYTGCTPVGYAMQPIETINVSABHNETHLFDIVSLKTPYITIV 738
QY 861 ATRVMAQVSLATVMLGGIAGFDDGSELRBSKFTSDSYSGDMFNRYDLQDNNTYV 920
Db 739 DERCVNHLATSIILSGIPEFHAGDELRSKSLDRDSYNSGDMFNRLDPSYNSNMGV 798
QY 921 GMPRIISDGSNEVITRVKEMVATPG-EAELKQMTAFYOELT---ELKRSPLFTLGDS 976
Db 799 GLPRPDHESNMPIL---KKRLANSPYKPKDKHITAVENFTNLQIRSSFLRLRSK 855
QY 977 AVMKRVDFRNTGSDQAGLLVMTVDGKAKA---SLDSRLDGLVVAIINAP-ESRTLN 1031
Db 856 DIEDVRFHNNVPSWIPGLIAMSIEDG-HAGAPGLSQIDPKFOYIVITINQPTETKFVN 914
QY 1032 -EFAETIOLSAIQOTAGENSLANGVQIAAGTVTLPAWSVAVLELPQ 1078
Db 915 PDLRAKSLQLHPVOSTSGDTVVKESKYEPTGCTTIPKSTAVFEPR 962

```

```

RESULT 3
081638 PRELIMINARY: PRT: 962 AA.
AC 081638:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE PULLULANASE-TYPE STARCH DEBRANCHING ENZYME.
GN ZPT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOSPERM.
RA Beatty M.K., Rahman A., Myers A.M., James M.G.;
RT "Genetic and biochemical characterization of ZPT1, a pullulanase-type
RT starch debranching enzyme from maize."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF080567; AAD11599.1; -
DR MENDEL: 32273; Zeama;24801;32273.
SQ SEQUENCE 962 AA; 106239 MW; 5F5AAB7C6EBD5D10 CRC64;

```

Query Match 25.4% Score 1432.5; DB 10; Length 962;  
 Best Local Similarity 35.7%; Pred. No. 1.4e-82;  
 Matches 336; Conservative 173; Mismatches 336; Indels 93; Gaps 24;

```

QY 181 GVALAE-----AHVWDKNTLMPGCGDKPIVRLYSSHSKVAADGEG-KFTDRYL 229
Db 73 GVAVAESAGFLDARAVYWTSLAMNISDKTSLFASNNATMKSSQDMKGGDSKV 132
QY 220 KLTPTT--VSQOVSMRFPHLSSYAFKLDPDANVDELQGTVAIAAEDGLISATQVQ 287
Db 133 ELQPENDEGLPSSVTOKFPFISSTYAFRIPSSVDVATLVYCOLAVASFDAHGKRDVTCIG 192
QY 288 TAGVDDAYAAEAALSTGAQIADGVTFRVAPTAQOVYVYVYASDAKKVISHPMTR-- 345
Db 193 LPGVLDDEFAITG---PLGTISECAVSMTYLAAPTAODVSSTFGP-----AGPLETV 244
QY 346 --DSASGAMSGOGSDLGAFYRYATVYHPOSKRVEQYEVTPYAHSLSTNSSEYQVVD 403
Db 245 QNNEINQWVSATGPRNWEKRYLLEVTVYHOTTGNIKCLADPAPARGLSANSTRMVLVD 304

```

```

QY 404 LNDALKPDGMDNLTPHAKTKAD-LAKMTIHESIIRLSAMDQTVAPELRGKIALTA 462
Db 305 INNETLKLPLAMGLA---AEKPRLDSSFSDISIEILHIDRFSADHSDIVDCPPRGGFCAPTF 361
QY 463 GSNVNOHKLTLASGVTVVELLPVFDLATVNEFSKQVADIQOPSRCEVSAVSSSEF 522
Db 362 QDSVGIETHLKLTLSDAGLTVHLLPSPFOGCGVD-----IKS--- 397
QY 523 ACYCSGSSTVEEV-LNQLKO-SDSODNPQVALNTLVLAQTDSYNMGYDPFHFTYPEGSYA 580
Db 398 -----NMKCVDELELSKLPPSGDLQ-----QAAIYAIQIEDYINMGYNVYMGVPRGSYA 447
QY 581 TDEPTTRIKERFTMIQAIKODLGMNVINDVYVNHNTAAGPDRTSYLDKIVPMWYQRLN 640
Db 448 SNPDGSRRIIEYRLVQALNR-LGLRVYMDVYVNHLSGSPRAITSVDLKIIPGYTLR-R 505
QY 641 ETTGSVESATCCSDSAPEHRMFAKLADSLAVTWDYKIDGRFPLDMLGHPRAQLISAME 700
Db 506 DSNQGTENSAAVNTASEHFNVDRLIVDDLWMAVYKYVDFRFDLMGIMKRTYMRAS 565
QY 701 RIKALNPD-----IYFEGEGMDSN---QSDRFELASQIMLKGTIGTFSDRLRDSYR 749
Db 566 ALQSILIDHGYDGSKITIYXGEMNFGVAAENORGLNSQLMKSSTGIGSFNDRIDALN 625
QY 750 GGGPDSGDALRQNOGIGSGAGVLPNELASLSDOYR---HLADTLRLGAGNLADFYM 805
Db 626 GCSPP--GNPQ--QOFSGLFLEPNGFYQGNDETRLRLATYVADHIGLGNLKDYYV 681
QY 806 IKDCAAKKSGEI-DYNGAPGCIYADPTTEVNVYVSKHDNQTLMDSIKSASOEDLATRV 864
Db 682 ISHTGEARRKSGEIRTFDGSPPVYASSPIETINVASAHDETFLEDISLTPMDLSIDERC 741
QY 865 RMQAVSIATVMLGCGIAGFOOGSELRBSKFTSDSYSGDMFNRYDLSQDNNTYVNM-P 923
Db 742 RINHLSTSIALSGIPEFHAGDELRSKSLDRDSYNSGDMFNRLDPSYNSNMGVGLP 801
QY 924 RISDGSNTEVITRVKEMVATPGEAELKQMTAFYOELTELKSSPLFTLGDSAVAKRYD 983
Db 802 REKNESGSLVPGVYVMSIEDARNDHDMQIDETFSQVAVFVNCPEVSIETPLASLRL 861
QY 984 FRNTGSDQAGLLVMTVDGKAKA---ASLDSRLDGLVVAIINAP--ESRTINEFAGEL 1038
Db 862 FRNTGSLVPGVYVMSIEDARNDHDMQIDETFSQVAVFVNCPEVSIETPLASLRL 921
QY 1039 QLSAIQOTAGENSLANGVQIAAGTVTLPAWSVAVLELPQ 1078
Db 922 QLPVQVNSSDALRQSAVDATGRTYVKKRAAVFEPR 961

```

```

RESULT 4
048541 PRELIMINARY: PRT: 904 AA.
AC 048541:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE LIMIT DEXTRINASE.
GN HVLN99.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. IGRI.
RA Lok F., Kristensen M., Planchot V., Leah R., Svendsen I., Svenson B.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF022725; AAD04189.1; -
DR MENDEL: 27102; Horvu;2480;27102.
DR INTERPRO: IPR000461; -
DR PFAM: PF00128; alpha-amylase; 1.
SQ SEQUENCE 904 AA; 99123 MW; C7116CC2DE99ED03 CRC64;

```





Query Match 24.6%; Score 1386.5; DB 10; Length 988;  
 Best Local Similarity 34.6%; Pred. No. 1.2e-79;  
 Matches 338; Conservative 164; Mismatches 359; Indels 117; Gaps 24;

```

QY 164 GNSAVYDSRADARFAAFVLAFAHVDKNTLMPGQDKPIRYLYSHSSKY-AADGEG 222
DB 64 GRGCVGECAAAVAASQGFVTDARAAYWTRSLIAMNVNDQSTSLFVASRDATMHVSGAI 123
QY 223 KFTDRYLKLP--TVSQOYSMRPHLSYAARFLPDNANVDELLOGETVAIAAEDGIL 280
DB 124 HGVDSKTELEPEHASLPDVAEKEFPFIRSTRFRVSSVVASLVCQLAVASYDAHGRH 183
QY 281 ISATOVOTAGVLDAYAEAAEALSYGAQLADGCVTFRVAPTAQOVDVYVYASDKKVTGS 340
DB 184 QDVTGLDLPGLVLDMEFAYTG--PLGAVFSDKDVLDYLAFTADQVAVCFYCDP-----A 235
QY 341 HPMTR-----DSAGAWMOCGSDLKGAFAFYRYAMTVYHPOSARKVEOYVTPRYAHSSTNS 396
DB 236 GPLLOTVQLKELNGVSVTPRYPRPENQYLYEVKVVHPSTSOVEKCLADDPYARGLSANG 295
QY 397 EYSQVVDLNDALSKPDGMDNLTPRHAOKTKADLAKMTIHESHIRDSAMPQVPAELRGK 456
DB 296 TRTMVNDVINESETLKPASWDELS--DEEPNLESFSDISITVELHTRDSAHNSTVDGNSRG 353
QY 457 YLAL-----TAGDSNMVOHLKTLTASGVTHVELLPVF 488
DB 354 FVHLHFRFLRLNLNDFCSPRTKHPGRIMETWQDSAGIRHLRKLKLSAAGLTVHLLPSF 413
QY 489 DLATV--NERSDKVADIQOQFSLRCEVNSAVKSEFAGYCDGSGTYBEVYNLKO--SDSQ 545
DB 414 HFASVVDNKSMMKFVD-----EAOIAKLKPPGSDQ 443
QY 546 DNPVOVALNTLVACTDSYNGCYDPFHITYPEGSAATDPEGTRIKERFTMQAIKODLGM 605
DB 444 -----QAIVYSIQOEDPYNMGYDPVLMGVPKGYASAPDPSRIETREKOVQALNR-IGL 497
QY 606 NVIMDVYVYHNTNAGPTDRTSVLDKIVPMYVYORLNETTGSVESATCCSDSAPHRMFAKL 665
DB 498 RYVMDVYVYHNTDSSGPGVSSVLDKIVPGYYLLRRN-VNGIENSAMNNTASEHFVYDRL 556
QY 666 IADSLAVWTTDYKIDGRFRLMGYHPRKAOLLSAMERIKALNPD-----ITFFGEGMD 717
DB 557 TVDDLLMAINIKVQDGRFLDMGHIMKSTIRAKSAIRSLTRVHGYYGKITYLGEQMD 616
QY 718 SN--OSDREIASQNLKGTGIGTFSDRLRDSYRGSGPPDSGDALRQNGIGSGAGVLP 774
DB 617 FGEVAAQHKRGINASQIMSGTIGISFNDRI RDYVNGCNP--GNPLQ--QGFSGLFLER 672
QY 775 NELASLSDQVHR---LADLTRLGMAGNLADVYMDKDAAKKGRF-DYNGAPGCIYAA 829
DB 673 NGYYQGENADTRRELATYADHIQIGLAGNLKDYVLRTHTEARKGSDITYEDFSDPYGYS 732
QY 830 DPEEVYVYKHHNDQTLMDMISYKASQEAADLATRVHQAVALATVMGOGIAFDQDQSEL 889
DB 733 SPVETINIVYSAHNDERTLPDIYSIKTRPGLSTIDEKCRINHLASSIALSQQIPFFHAGDEI 792
QY 890 LRKSKFTSDSYSGDMENRVDYSLQDNNYVNGMPRISDDGSNEVITRVKEMATPGEAE 949
DB 793 LRKSLDRDSDYSGDMENKIDFTYETNNMGVGLPPRDKNEENMLI---KPRLENPSFRP 849
QY 950 LKO--WNAF--YQELTEBLKSSPLFTLIGDSAAWKRYDFKNTGSDQAAGLLVMTVDGKM 1005
DB 850 LKHHIISAFNFQDLKIRSSPLFRLSTASDIEQRYRFRNHTGSMVPGVYVMSIKQAON 909
QY 1006 AG---ASLDRSLQDLVAIAAEP--SRTLNEFAGETLQSAIQOTAGESLNGVQIAA 1060
DB 910 EKCKMAOLDNKNPSVYVYIIFWVCEHVSIEITHLASLCLLHPIQVNSDALVYQSAEAS 969
QY 1061 DGTVTLPAWVAVLELPO 1078
DB 970 KCRFTVYPRRTTAVFVQPR 987

```

```

RESULT 8
ID P93416 PRELIMINARY; PRT: 986 AA.
AC P93416;
DT 01-MAY-1997 (TEMBLrel. 03, Created)
DT 01-MAY-1997 (TEMBLrel. 03, last sequence update)
DE 01-MAY-2000 (TEMBLrel. 13, last annotation update)
DE STARCH DEBRANCHING ENZYME PRECURSOR.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. FUJIKIRARI; TISSUE=ENDOSPERM;
RX MEDLINE=96275956; PubMed=8680309;
RA Nakamura Y., Unemoto T., Ogata N., Kuboki Y., Yano M., Sasaki T.;
RT "Starch debranching enzyme (R-enzyme or pullulanase) from developing
RT rice endosperm: purification, cDNA and chromosomal localization of the
RT gene."
RL Planta 199; 209-218(1996).
DR EMBL: D56002; BAA09167.1;
DR MENDEL: 16402; Oryza; 16402.
DR INTERPRO: IPR000461;
DR PFAM: PF00128; alpha-amylase; 1.
KW Signal.
FT SIGNAL.
FT CHAIN. 75 986 POTENTIAL.
FT SEQUENCE 986 AA; 109592 MW; FBA899E1634CAD27 CRC64;

```

Query Match 23.5%; Score 1326.5; DB 10; Length 986;  
 Best Local Similarity 33.8%; Pred. No. 7.8e-76;  
 Matches 333; Conservative 160; Mismatches 360; Indels 131; Gaps 26;

```

QY 164 GNSAVYDSRADARFAAFVLAFAHVDKNTLMPGQDKPIRYLYSHSSKY-AADGEG 222
DB 64 GRGCVGECAAAVAASQGFVTDARAAYWTRSLIAMNVNDQSTSLFVASRDATMHVSGAI 123
QY 223 KFTDRYLKLP--TVSQOYSMRPHLSYAARFLPDNANVDELLOGETVAIAAEDGIL 280
DB 124 HGVDSKTELEPEHASLPDVAEKEFPFIRSTRFRVSSVVASLVCQLAVASYDAHGRH 183
QY 281 ISATOVOTAGVLDAYAEAAEALSYGAQLADGCVTFRVAPTAQOVDVYVYASDKKVTGS 340
DB 184 QDVTGLDLPGLVLDMEFAYTG--PLGAVFSDKDVLDYLAFTADQVAVCFYCDP-----A 235
QY 341 HPMTR-----DSAGAWMOCGSDLKGAFAFYRYAMTVYHPOSARKVEOYVTPRYAHSSTNS 396
DB 236 GPLLOTVQLKELNGVSVTPRYPRPENQYLYEVKVVHPSTSOVEKCLADDPYARGLSANG 295
QY 397 EYSQVVDLNDALSKPDGMDNLTPRHAOKTKADLAKMTIHESHIRDSAMPQVPAELRGK 456
DB 296 TRTMVNDVINESETLKPASWDELS--DEEPNLESFSDISITVELHTRDSAHNSTVDGNSRG 353
QY 457 YLAL-----TAGDSNMVOHLKTLTASGVTHVELLPVF 488
DB 354 FVHLHFRFLRLNLNDFCSPRTKHPGRIMETWQDSAGIRHLRKLKLSAAGLTVHLLPSF 413
QY 489 DLATV--NERSDKVADIQOQFSLRCEVNSAVKSEFAGYCDGSGTYBEVYNLKO--SDSQ 545
DB 414 HFASVVDNKSMMKFVD-----EAOIAKLKPPGSDQ 443
QY 546 DNPVOVALNTLVACTDSYNGCYDPFHITYPEGSAATDPEGTRIKERFTMQAIKODLGM 605
DB 444 -----QAIVYSIQOEDPYNMGYDPVLMGVPKGYASAPDPSRIETREKOVQALNR-IGL 497
QY 606 NVIMDVYVYHNTNAGPTDRTSVLDKIVPMYVYORLNETTGSVESATCCSDSAPHRMFAKL 665
DB 498 RYVMDVYVYHNTDSSGPGVSSVLDKIVPGYYLLRRN-VNGIENSAMNNTASEHFVYDRL 556
QY 666 IADSLAVWTTDYKIDGRFRLMGYHPRKAOLLSAMERIKALNPD-----ITFFGEGMD 717

```

[illegible]

Query Match 12.7% Score 716; DB 2; Length 849;  
Best Local Similarity 24.7%; Pred. NO. 4.1e-37;  
Matches 243; Conservative 139; Mismatches 315; Indels 286; Gaps 35;

```
QY 54 AVTAVENQAVIHLVDIAGTSSADYSSKNTLYLNNFECALISAPVADMDVSTTPSGS 113
Db 24 SISFETELIHYHRWDG-----NYDGMNLMW----- 51
QY 114 DRYGPWVPLPKKESGICINVIYRDG-----TDKLIDSLRVAFGFDTBRTVS----- 160
Db 52 -----WPEPISIEGAAVYFTEEDPGVAAKAFPELTKVGLIYVAGREVDVAMDRP 105
QY 161 -----VIAGNSAVYDSRADAFRAAFGVALAEAHMVDKNTLLMPGODKPIYRLV 209
Db 106 IRIKOGAEVWLLOGIEQYITTKPDT-----SPRVLQFQARQYITENAV 149
QY 210 YSH-----SSKVAADGCKFTDRILKTPPTVSGOVSKRPHLSSTYAA--FKLPD-NA 259
Db 150 LYGVDVTVKAKVYDQGPFLKARERKANPDIS-----RTHNVVYVLAEPKIDEDVK 204
QY 260 NYDELLOGEFTVIAAEDGILISATOVATGYLDDAYAEAAEALSGAOLADGVTFERYM 319
Db 205 DVQVELEGKPA-----RVIMMELIDKITYDGPLOGFEYTPPE---KTTIRW 247
QY 320 APTAQOVVVVYS--ADKKVISHPMTSDASGAMSGSDLKGAFYRYAMTVYHPOS 377
Db 248 SPVSKTVDLVLYKNNMDCKEPTGVVPM-KYINGAMEAVLEGMMDFYKIRFESYG---- 302
QY 378 KVEQVEVTPPYAHSLSSTNEYSQVVDLNDLSALKPDGMDMLTPHQAOKTADLAKMTIHES 437
Db 303 --EYRESVDYFSKAAVTKNAKSAIIDF--SKTNPESEMEKVARPPLVAPDAI---IYEI 354
QY 438 HIRDSAMPQVPAELRGKYLAL----TAGDSNMVOHLKTLASGVTYHELLPVPDLATV 493
Db 355 HIADMTGLDMS-CVKKKALYLGITEKGTGPNQVTTGLDHLVELGYTHVHILPMFPWGTG 413
QY 494 NEFSKRVADIQOPFSLCEVNSAVKSEFAGYCDGSJVEEVLNQLKQSDODNPQVAL 553
Db 414 DE-ADK-----DEFER----- 422
QY 554 NPLVATQDSYNNGYDPEHTYPEGSYATDP-EGTTRIKEFRMIOAKIKODLGNVIMDVY 612
Db 423 -----SYNNGYDPEHTYPEGSYATDPINPYTRILEVKQVKAHLHN-GIRVILDAV 473
QY 613 YNITNAAGPTDRTSYLDKIVPMYQORLNETTGSVESATCCSDSAPHRMAKLIADSLAV 672
Db 474 PHTMGVGV--MSPPDAQVPIYFYRIDKTGAVLINESGCGVNASRPMKRYIVDTLKW 530
QY 673 WTTDIDYIDGFRDLMGYNHRAOILSAMERIKALNPDIYFFEGEGWDS--NOSDFEIASQIN 731
Db 531 WYTERYKIDGFRDQMGIMKVTMLAIKSELSKIEPVSVLGYEPMGGMGAPIRF---GKAD 587
QY 732 LMGTOGTCSDRLRDSRGCGPPDSGDALRQNGIOSGACVLPNELASLSDDQVRHLADL 791
Db 588 VGGTGTAAENDEFROALRG-----SVFNATVKGf--- 616
QY 792 TRIGMAGNLADFAVMDKDAANKG--SEIDYNGAPGYAADPTEVYVYNSKHDNQLMDM 849
Db 617 ----LMGALA-----KELGVKRCVAGSIEYDEVIRSFPAKDPQETIINYEYVHNHLMK 666
QY 850 ISYKASQEDLATRV-----MOAVSLATVYMLGOGIAPDOGSSELLRSKFTDSDYD 901
Db 667 -NYLAAO-AD--TYVWTEEMLKDAOKLAGAILITSGOIPFLIAGODFARTKKEFDENSYK 722
QY 902 SCDMFRVYVSIQDNNYVNGMPRIISDGSNYEYITVKEMVATPGAEALUKOMAFQOELT 961
Db 723 SPISINGLDYA-----RAAEFTIDVFNYYKGLI 749
QY 962 ELKSSPLFTLGGSAVMKRVDF 984
Db 750 ELKSHIAFRORATIEDIRKKLTF 772
RESULT 11
ID P70983 PRELIMINARY: PRT: 1938 AA.
AC P70983;
```

```
DT 01-FEB-1997 (Tremblrel, 02, Created)
DT 01-FEB-1997 (Tremblrel, 02, Last sequence update)
DT 01-OCT-2000 (Tremblrel, 15, Last annotation update)
DE ALKALINE AMYLOPULLULANASE.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSM-1378;
RX MEDLINE=96394539; PubMed=8798645;
RA Hatada Y., Igarashi K., Ozaki K., Ara K., Hitomi J., Kobayashi T.,
RA Kawai S., Watabe T., Ito S.;
RT "Amino acid sequence and molecular structure of an alkaline
RT amylopullulanase from Bacillus that hydrolyzes alpha-1,4 and alpha-1,6
RT linkages in polysaccharides at different active sites.";
RL J. Biol. Chem. 271:24075-24083(1996).
DR EMBL; D78258; BA011332.1; -.
DR HSSP; P06278; IVS3.
DR INTERPRO: IPR000461; -.
DR INTERPRO: IPR000834; -.
DR INTERPRO: IPR001899; -.
DR PFAM; PF00128; alpha-amyase; 2.
DR PFAM; PF00746; Gram_pos_anchor; 1.
DR PROSITE; PS00133; CARBOXYPEPT_2N_2; UNKNOWN_1.
SQ SEQUENCE 1938 AA: 214917 MW: A29D3F715EB6A72E CRC64:

Query Match 12.3%; Score 692.5; DB 2; Length 1938;
Best local Similarity 24.2%; Pred. No. 5,1e-35;
Matches 281; Conservative 168; Mismatches 387; Indels 325; Gaps 51;

QY 11 LGSVLV--LSGCDNGSSSSSSSGNPTPDQDVVRLPDAVGEAVTAENOAVIHLVD 68
Db 895 IGLVLNTINGDKDGDGKAVELFSPDL-NEIWIKGSDEVLFEYVDLPANVRIHY-- 950
QY 69 IAGTSSAADVSSKULYLMNNFECALISAPVADMDVSTTPSGSDKYGPVWVPLNKES 128
Db 951 -----ERTNADYEGMGLMNNED-----VESPSDGMWPGADADAIGIKYGYDIKLEDA 1000
QY 129 GCINIVVROGTDLIDSLDLRVARGDFTDRTVSYIAGNSAVYDSRAAFRAFPVALAEAH 188
Db 1001 NKIGLFLVN-----KQSGGOTGDMTFMLK----- 1025
QY 189 WDKNTLLMPGGODKPIVRLVYS-----HSSKYAADGEGKFTDRYLKTPTVSQOVSMR 243
Db 1026 --QYNGLFYKEGEDKYVTNPYGVPLALVSGEVLSD-----KLISLFTFRREGLD 1073
QY 244 FPHLSYAAFKLPDMANV--DELOGEYVAIAAE--DGILISATOV-QTAGV----- 291
Db 1074 LEELEKOLEIKVDGNDVSEFTDVTIEGKTVHVGGEFDELEKIPFSYTYLDRITSVKSGWK 1133
QY 292 -LDDAVAEAAEALSYAGN-AQDGVTFRVWAPTAQOVVVVYSA--DKYIGSHPTMRDS 347
Db 1134 LIDEMKAYDGK--LDAELHEDGTATLKWSPKADNVSVLKYQVQONEVYDITENAKGD 1190
QY 348 ASGAMSMQ-----GGSDLKGAERYAMTVYHPOSKVEQYEYV--DPYASHLST-NS 396
Db 1191 -RGVMSVKLTKMDWTGLSLKGYHYHEIT-----HGDVTNLALDPYAKSMAAWN 1239
QY 397 E-----YSQVVDLNDLSALKPDGMDMLTPHQAOKTADLAKMTIHESHIRDSAMPQYVP 450
Db 1240 EAGDKVKAALVDI--GSIGPE-LDYADIPGFEKREDTI--IYEVHVDFTS-DPNIG 1291
QY 451 AELRGKYIALITAGDSNMVQHLKTLASGVTYVHLLPVPDLATVNEFSKVDKADIQOPFSL 510
Db 1292 EDLKAQFGTF--ASVVEKLADYIQELGYTHIQLLPVMSTYFSNER-----ESGERM 1339
QY 511 CEVNSAVKSEFAGYCDGSJVEEVLNQLKQSDODNPQVALNTLTVAGTDSYNNGYDPF 570
Db 1340 LE-----YASTGT-----NTNMCGIDPH 1356
```



QY	571	HYTPESVYATDEEG--TTRIKERTIQA1-KODGMMNYIMYVNHNNAGSPDRSYL	628
Db	1357	NYEELSMTSENEDEPLRIKKEKMLINEIKRDMG--VLDYVNH-----TAQVHF	1408
QY	629	DKIVPMVYQRIANETTSVSBSATCCSDSAPEHMFALADSIAVTTTYKIDGFRFDLMG	688
Db	1409	EDLVPNTYHFM-DADGPRTSFGCRGLGTTHEKMSRRVLVDSJKHWYDEKVDGFFEDMG	1467
QY	689	YHPKAOILSMEKRIKALNPDIYFFGGCV--DSNOSDRFIASQINLKG7-GIGTFSPRL	744
Db	1468	DHDESSIQIAFDEAKKINPNINWIGGWNTPAGDECEPVQADDOOMQOTEAVGSFSDFF	1527
QY	745	RDSYRGGPPDSGDLKRONQIGSGAGVLPNELASIDDOYVRLHADLTRLGMAGLADFY	804
Db	1528	RNEIK-----SGFG-----SEQPPFIT-----	1545
QY	805	MIDKGAKKKGEI--DYNGAPGVYAD-PTEVYNYVSKHNDQTLMDIMYSKASGEADLA	861
Db	1546	-----GGAVNVQOIFDNIRKAPHNFMADPGDVGYIEAHNMLTYIAQSIKKRPEIA	1600
QY	862	-----TRYRMAVSLATVMLOGQIAFPQSGELLKRS-----	894
Db	1601	ENDELHKRIR--VGNAMVLISQGAFFLAGGEFRTQKMPARATEARYSTWTDAG	1657
QY	895	-----FTDSYSGMGNFRVDYSLODNNYNVGMERISDDSGNSVYTRVKEWATPGE	947
Db	1658	NPEYFPFHDSYDSSDILNRFDMK-----ATDAEKYPVNNVTRDYTA-----	1701
QY	948	AELKOMATFVQOELTELKSSPLTLCDSGVYKRVDFRTGSDQAGILVMTVDGCKMAG	1007
Db	1702	-----GLEELKRRSDAFRLRSRELYDSNTYVAPAEKEDLV-----AY	1742
QY	1008	ASLDSRLDGLVVAINAPESRTINERAGETLQASLQVAGNSLAN-----GYOIAADG	1065
Db	1743	RSVTAAGVETTYTNMATSRTIL7-LGODITGCVYVVAEEANVAGVAPAGFELTAG	1800
QY	1063	TVTLPAMSVAVLELPOGEAG 1083	
Db	1801	-TLEPLTTVVVRV-GEQEG 1818	
RESULT	12		
ID	Q59319	PRELIMINARY;	PRT; 825 AA.
AC	Q59319;		
DT	01-NOV-1996 (TEMBLrel. 01, Created)		
DT	01-NOV-1996 (TEMBLrel. 01, Last sequence update)		
DT	01-MAY-2000 (TEMBLrel. 13, Last annotation update)		
DE	ALPHA-DEXTRIN 6-GLUCANOHYDROLASE (EC 3.2.1.41)		
DE	ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE (PULULANASE)		
DE	(PULULAN 6-GLUCANOHYDROLASE) (LIMIT DEXTRINASE) (DEBRANCHING ENZYME)		
GN	(AMYLPECTIN 6-GLUCANOHYDROLASE).		
OS	Caldocellum saccharolyticum (Caldicellulositriptor saccharolyticus).		
OC	Bacteriia, Firmicutes; Bacillus/Clostridium group;		
OC	Thermocaneobacter group; Caldicellulositriptor.		
OX	NCBI_TaxID=44001;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Albertson G.D., McHale R., Gibbs M.D., Bergquist P.L.;		
RL	Eur. J. Biochem. 0:0-0(0).		
CC	-1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYSES (1-6)-		
CC	ALPHA-GLUCOSIDIC LINKAGES IN PULULAN AND STARCH TO FORM		
CC	MAITOTRIOS.		
DR	EMBL: L33876; AAB06264.1; -		
DR	INTERPRO: IPR000461; -		
DR	PFAM: PF00128; alpha-amylase; 1.		
KM	Hydrolase; glycosidase.		
SC	SEQUENCE 825 AA; 95732 MW; 92C2082E9B35F5AD CRC64;		

Query Match	12.28;	Score 685;	DB 2;	Length 825;
Best Local Similarity	28.28;	Pred. NO. 3.6e-35;		

	Matches	217;	Conservative	100;	Mismatches	242;	Indels	210;	Gaps	28;													
OY	296	YAFAEALS-----	YGADLGDGVTFFRYMAFPAQOVDVVVYSADKIVGISHPMTRD							346													
Db	191	YARPREILITRNEFYAGELGARPEPYGTIFRLMAPTAVKVIQLPDEHENFKFEKMER-								249													
OY	347	SASGAMSGGSDDLKGATYYRAAMY -HPQRKEQIYEVTTPYAHSLSTNSEYSQVDLNL								405													
Db	250	AENGTMWDYLLPLDKLNHYLYEIMYINVEDEGYLTVVAVPPDYSAASSNSNGKSYIPDPA								309													
OY	406	DSALKPDGM-----DNLTMPHAKTKRADKLAKMTIHESHINDLSAMDQT--VPALRGKY								457													
Db	310	DGLT--DGWQADSFYDNI-----EKDDA--ITYENMHVDFTI--DOSSGIGENLRKF								357													
OY	458	LAL-----TAGDSNMVOHLKTLASGYTHVELLPVFPLATVNFESDRKVADIQQPFHRIC								511													
Db	358	LGFCQEGEYKKGGISITGLHLKEL---GVTHNHLPIPSFGSYD-----								398													
OY	512	EYNSAVKSSERFAGYCDGSTVEVLNLQKSDODNPVOALNTLVAOOTSDYNMGIDPH								571													
Db	399	-----KNDPKR-----YMWGYDEVL								413													
OY	572	YTFPGSYATDPEGTTIRIKFEPTMIOAIKODLGNVINDDVYNHT--NAGSPTRDSYLDK								630													
Db	414	YTCEPYMYSTSGGLEALKELRTMKKHEN-GIVADVDFVNHTIYHKG--GKSISTDK								470													
OY	631	IVPWYQHILNETTGVSESAT--CCSDSAPENHFAPKLIADSLVAWTTDYKIDGFREDLNGY								689													
Db	471	IYPEFYHVDD-YGDYSNATGGCNELAPEKRVKFIIDTIITYWEERHIDGFRFDLMGL								529													
OY	650	HPKAOLISAMERIKALNDIYFPFGGWDSNOSDRF--BIASQINL--KGTGIGTFSRLR								745													
Db	530	LTVKOSROVANEVRKRKPALUYEGGWYMGNSTCLVEBEMATILLSCHOGYSIGLENDRIK								589													
OY	746	DSVRGSGPFDSDALRKQNOIGSGAGVLPNELASLSDQVRIHALDTLGMAGNAIADRYM								805													
Db	550	DAIRG-----DLDGKTYGT-----VHGNLSDV--								611													
OY	806	IDKDGAARKSEIDVNGARGCYAAPTEVVVNVSKHDNOTLDMISYKASOE---ADLA								861													
Db	612	----GRLNQG-----IKAAIDPRAKEPDCVNVSGHDNILTFDKIQKMVGEDIIDWIRA								663													
OY	862	TTRYMQAVSLATVMLGCGIARDQGSSELLRSKSFTRDSYDSGDWFNRVYDSLQDNNYNG								921													
Db	664	TR-----LANAIVLTSQCVAFILHGCVEFNRSGGHPRNTYANAGDINKIDWSLKEKFFDT								717													
OY	922	MPRIDDSNENVITRKNEMATPEAECLKQTAAYOELTELKRSPLFTLDDGSAVVKR								981													
Db	718	-----FKFYCDLIRNRHNIAFRMSGGEIKKY								745													
OY	982	VDFRNMGSDOAGLVMTVDDCMKAGASLDSRLDLVAINAAPSRFL								1030													
Db	746	LRFIAPRCGVAAFITTYPYDENKK-----IIYAYPNFKKKIL								783													
<hr/>																							
RESULT	13																						
ID	034587	PRELIMINARY:	PRT:	718	AA.																		
AC	034587:																						
DT	01-JAN-1998	(TREMBLrel. 05, Created)																					
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)																					
DT	01-MAY-2000	(TREMBLrel. 13, last annotation update)																					
DE	AMTX	PROTEIN.																					
GN	AMTX.																						
OS	Bacteria; subtilis.																						
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;																						
OC	Bacillus/Staphylococcus group; Bacillus.																						
OX	NCBI_TaxId=1423;																						
RN	[1]																						
RP	SEQUENCE FROM N.A.																						
RA	Lapidus A., Galleron N., Sorokin A., Ehlich D.,																						
RL	Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.																						

RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
MEDLINE-98044033; PubMed-9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertolo M.C., Bessieres P., Bolotin A., Borcherdt S.,  
RA Boursier R., Boursier L., Brans A., Braun N., Brignell S.C., Bron S.,  
RA Brouillette S., Brusch L., Caldwell B., Capiano N.J., Carter N.M.,  
RA Chol S.K., Codani J.U., Conerton I.F., Cummings V.J., Daniel R.A.,  
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entlin K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,  
RA Gilm S.Y., Glaser P., Goffeau A., Goughly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henalt A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullio M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaber-Blanchard M., Klein C.,  
RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mouel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schoeter R., Scoffone F.,  
RA Sekowska J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Takashashi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viart A., Wambolt R., Wedler E., Wedler M., Weltenegeger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
RT \*The complete genome sequence of the gram-positive bacterium Bacillus  
RT subdilis\*;  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,  
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AF008220; AAC00283.1;  
DR EMBL: Z99119; CAB14971.1;  
DR INTERPRO: IPR000461;  
DR PFAM: PF00128; alpha-amylase; 1.  
SO SEQUENCE 718 AA; 81104 MW; 96f9eb767877e93 CRC64;

Query Match 11.5%; Score 647.5; DB 2; Length 718;  
Best Local Similarity 27.1%; Pred. No. 6.9e-33;  
Matches 224; Conservative 99; Mismatches 250; Indels 253; Gaps 33;

OY 286 VOTAGVLDAAAYAAALSLGADLADGCVTFKRWAPPAQOVVYVSADKKVIGSHF--- 342  
DB 91 IRTAAADDEFFYVG---ELGAVYTADHVFVKWAPPAASAALKL-----SHRNKS 137  
OY 343 -----MTPDSAGSAMWGSGSDLKGAFFRYAMTVYHQSKVQEYETOPAYSLSTNSE 397  
DB 138 GRFFQMR-LEKGVYAVTVYGDILHG--YETLPGT-----CNSEMETVVDQYAAAVTNGE 190  
OY 398 YSOGVVDLNSALKPD--GWDNLTMPAOKTKADLAKMTIHESHIRLDSAMDQTVPAELRG 455  
DB 191 KGVY-----LRPDQMKWAPLAKPFSHPYDA-----VIYETHLRDPSIENSGMIN-KG 237  
OY 456 KYALALTA-----GDSNWKQHLKTLASAGVTVYVELLPVDLTVNVEFSKAVADIQOQFS 508  
DB 238 KYALLETDTOTYANSSSSGLAAYVKEL--GVTFVELLPVNDFAGVDE-----EKPL- 285  
OY 509 RLCEVNSAVKSEFAGYCDGSGTVEEVLNQDKSDSDQDNQVQALMTLVNQSTYKMGYD 568  
DB 286 -----DAYMNGVN 293  
OY 569 PPHYTVPEGSYATDP-EGTTRIKEFTMIQAIKQDLGMNVIMDVYVNHGMAAPTDTSTY 627  
DB 294 PLHFAPBEGSYASNPDPQTRKTELKOMINTLNGH-GLRVLIDLVVFNHYKR-----ENSP 348

OY 628 LDKIWPYQORLNETTGSVESATCCSDSAPEHRMFAKLADSLAWTTDYKIDREFDL 687  
DB 349 FEKTVPCYFPHRDECGMPSNMGVGNDIASERRNARFIADCVYVLEEVNVDGFRDL 408  
OY 688 GYHKAOLISAMERIKALNDIYFFGGWD-SNDSREFIASQINTL-KGIGIGFSRRLR 745  
DB 409 GILDIDIVLWKEKATKAKPGILLFFGEGWDLATPLPHEOKAALANAPRMGIGFNDMFR 466  
OY 746 DSVGGGPPD--SGDALRONOGIGSGAGVLPNELASIDQVYHNLADLTRLGAGN--- 799  
DB 469 DAVK-GVTFILKATGFL-----GNE-----SAQAVH-----GAGSSGW 504  
OY 800 --LADFYMDKGAAGKSGEIDYNGAGVAADEFEVYVYVSKHNDOTLMDIYSKASQ 857  
DB 505 KALAPIV-----PESQGINVESHDNTPMDKNSFALPQE 540  
OY 858 ADLATRPMQAVSLATVYMLGGLAFDQGSSELRKSFRTSDSGWPNRVDSLODNN 917  
DB 541 NDSRRSR-ORLAVALIILLAGVFFHSGOEFRTKGVNSYSSDSINQDLDRRET- 598  
OY 918 YVGMPIRISDGSNYEYITRYKEMVATPGEAELOMTAFYQELTELKSSPLFTLGSA 977  
DB 599 -----FKEDVHYIRLISLKKHAPFRL----- 621  
OY 978 VMKRVDFRNTGSDQAGLVMTYDDGKMGASLDSRLD-----GLVVAINAAPES---- 1027  
DB 622 -----RSADIQRHLECIITLKEHLIAVRLVD--LDEVDEMKDIIIVIHASPSDSEWR 671  
OY 1028 -----RTLNEFAGELQLSAIQOTGENSLANGVIAADGYTL 1066  
DB 672 LPNDIPYRLCLDPSGFQEDPTEIKTVAVNGI-----GVYIL 708

RESULT 14  
ID 069008 PRELIMINARY; PRT: 718 AA.  
AC 069008;  
DT 01-AUG-1998 (Tremblrel, 07, Created)  
DT 01-AUG-1998 (Tremblrel, 07, Last sequence update)  
DT 01-MAY-2000 (Tremblrel, 13, Last annotation update)  
DE PULULANSE.  
GN PUL.  
OS Thermus sp.  
OC Bacteria; Thermus/Delinococcus group; Thermus group; Thermus.  
OX NCBI\_TaxID=275;  
RN [1]  
RC STRAIN-IM6501;  
RA Kim T.J., Kim J.W., Park K.H.;  
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF060205; AAC15073.1;  
DR INTERPRO: IPR000461;  
DR PFAM: PF00128; alpha-amylase; 1.  
SO SEQUENCE 718 AA; 80446 MW; 8f7c0d22405eab8 CRC64;

Query Match 11.2%; Score 630; DB 2; Length 718;  
Best Local Similarity 27.1%; Pred. No. 9e-32;  
Matches 232; Conservative 92; Mismatches 282; Indels 250; Gaps 35;

OY 150 ARGDFTDRIVSYTAGNSAVYDSRADAFRAAFGVALAEAHVNDKTLIMPGGDKPIVRLY 209  
DB 10 AYIDEMQOIV-VLAPKSLGFGMAFP-----TIVAPSGGEIP--LS 47  
OY 210 YSHSKVAADBEKFTDRYKLTPTTYSQOVSMSRFPULSYAFAKLPDMANVELLOGET 269  
DB 48 VQHEVDY--GE--TVKYV-----CRFA-----SAFER-----GAT 73  
OY 270 VAIAMAEGLISATQVOTGVL-----DDAYAAEALSYGOLADGCVTFRYMAPTAQ 324  
DB 74 YWVRSCRG-----EETDVQIAGAVYRTAFDFRFFYDG--PLGAEYILKEQVTFRYMAPTAT 126

```

QY 325 QVDVYVSADKKVYIGSHPMTRDSASGAMSGSDLKGAFTRYAMTYVHPQSRKVEQYEV 384
DB 127 AVSKVLPHPLDELRCVPLVR-GERGVSAVYPGDMERARITYIACI-----NRYMR-EA 179
QY 385 TDPAHSLSTNSEYSOVVDLNDLSALKPDGMDLMPH-AQTKRADLAKMTIHESHIRPLS 443
DB 180 VDPATATAVSNGEFGVYIDMEKTKLAP---PSLPLPLCLSTDA-----IIEISIRKFT 231
QY 444 AMDQTPAEELGKYLAL-----TAGDSNMVQHLKTLASGVTHVELLPVFDLATVNEFSK 499
DB 232 SHPDS-GAVHGXKYLGAETNTSGPNCATGLSYVKELGYVHYDLMPEMDPAGVDE----- 286
QY 500 VADIQPFPSRLCEVNSAVKSEFAGYCDSGSTVEEVLNQLKQSDQDNPOVALNTLVAQ 559
DB 287 -----RDPQA----- 291
QY 560 TDSYNGDPPEHYVTPGGSYATDP-EGTTRIKERFTMIQAIKODLGMVIMDVYVNHNA 618
DB 292 --ATNMGYNPLHLAPBESATADPADPYARKLVEIKQALHTLHEN-GLKVVMDAVYVNHV-- 346
QY 619 AGPTDR-TSVLDKIVPYVYORLNETGVSSESATCCSDSAPHRMFAPKLADSLAVMTTDY 677
DB 347 --YDREGSPLEKLVPGYFRYDAYGQFANGTGVNDIASERBARMIYDSVVFMAKEY 403
QY 678 KIDPFREDLNGYHPRKQOILSAMERIKALNPDIYFEGSGMD-----SNQSRFEIASQIN 731
DB 404 GIDPFREDLNGVHDIEITKAVARDALDAIDPSILVYEGSGMDIPTPLPPEQKATMANAKOL 462
QY 732 LKGTGIGTFSRLDSVYGGGPF---DSGDALRONOGIGSGAGVLPNRLASLSDQVYVHL 788
DB 463 ---PRAFVFNDRFDAYK-GSTFHLPRGFAL-----GNPGG----- 495
QY 789 ADLTRLGAGNLMADPVMIDKGAAGKSEIDYNGAPGAYADPTEVNVYVSKHNDQTLMD 848
DB 496 REGYKLIAGSLR-----ALGSLFCHPQSIINIVYVCHDHHTFMD 534
QY 849 MISTKASOQADLATRYMNOAVSLATVMIGGIAFDQSGSEILRSKSTRTSDYSGDMFNR 908
DB 535 KMEANHDEPEMLRRKR-QKLTATVILAOIGIPLHSGOEFYRTKGSGNGSNYRSPDAVNO 593
QY 909 VDYSLQDNVYVWGPRLSDGNSNEVITRVKENVATPEALKQMTAVYOELTELKRSKP 968
DB 594 LDMERKS-----RIEDD-----VRVQGLIALKRAG 620
QY 969 LETLGDSAVMKRYDF 984
DB 621 AFRILATEAEVLRHFTF 636

RESULT 15
Q9K6N1 PRELIMINARY: PRT: 1072 AA.
AC Q9K6N1:
DT 01-OCT-2000 (Tremblrel.15, Created)
DT 01-OCT-2000 (Tremblrel.15, Last sequence update)
DT 01-OCT-2000 (Tremblrel.15, Last annotation update)
DE ALKALINE AMYLLOPOLIDIANASE.
GN BH3697.
OS Bacillus halodurans.
OC Bacteria: Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001519; BAB07416.1;
SQ SEQUENCE 1072 AA: 120674 MW: 5264235CF24619D9 CRC64:

```

```

Matches 262; Conservative 155; Mismatches 362; Indels 347; Gaps 48;
QY 7 NALVGLSVLLSGDNGSSSSSGNPDTPNQDVVLPVAVGAVTAVENQAVIHL 66
DB 96 NAENGFVLV-----NTATGDKDGN-KHQQLLSPEH-----KEIWKREGSDVFL 140
QY 67 ---VDIAGITSSAADYSSKILYLMNNETCALSAVADWMDVSTPPSGSD-----K 115
DB 141 YEPVDLPQHTVRIHERSDQSYHHWGLTWMKDVASPEEH-----PAGANPFSEQTSR 194
QY 116 YGPVWVITLNNESGGINIV--ROCTDKLISDLRVAAGDFTDRVSYIAGNSAVYDSRA 173
DB 195 YGAVYDELADANEIGFLVVRNENGDK-----DDDKTFADLANHPOLFIRQ- 242
QY 174 DAFRAFGVALAEAHVMDKNTLLMPGQDKPIVRLYSHSSKVAADGEGKFTDRKLTP 233
DB 243 -----GDDIYVNTNYPTIEGLVSGELISLTEMELRTS 276
QY 234 T-----TVSQOVSMPREPLSSYAEPKLPNANVDELLOGETVAIAAEDGI 279
DB 277 TIMLSEDDLAQIHLTDNGASVAFDHT-----IHDEKTV--ILHGSFDELAPYHYT 328
QY 280 LISATQVQTAG--VLDAYAAEAELSYGAOL-ADGVTFRVMAPTAQOVVYVVSAAK 336
DB 329 PADKTIVARACWRLIDALYGDGE--LGAVLHEDGSSLTTFWSPFLADHVSVYLDKDDQ 385
QY 337 ---VIGSHPMTRDSASGAMSW-----OGSDLKGAFTRYAMTYVHPQSRKVEQYEV-- 385
DB 386 YRIIIDDIDMTKGD-QGVWQVTLISENTGLNLAGYVHY-----KIERGDTKL 434
QY 386 --DPAVHSLST-NSEYSOVVDLNDLSALKPDGMDLMPHQAQTKRADLAKM-----T 433
DB 435 GLDPYAKMAAMNNGYSI-----GKAALVDPSSIGPELDFEIDGFEKREDAI 484
QY 434 IHESHIRLSAMDQVVPDELKGYLALTAGDSNMVQHLKTLASGVTHVELLPVFDLATV 493
DB 485 IYEVHVRDFTS-DPQIEBELTAQEGTF---ASFYDKLDYLEDLQTHIOLLPVMSY--- 536
QY 494 NEFSQVADIQPFPSRLCEVNSAVKSEFAGYCDSGSTVEEVLNQLKQSDQDNPOVAL 553
DB 537 -YMOGELANHERELA-----YSSSG----- 555
QY 554 NTLVAQDTSYNGVDPPEHYVTPGGSYATDPDG--TTRIKERFTMIQAIKODLGMVIMDVY 612
DB 556 -----NNYMGVDPHSYFSLSGMYSENPEDELRIKEFKHLIDEIRH-RMGVILDVY 607
QY 613 YNHTMAAPRTDRTSVLDKIVPYVYORLNETGVSVSATCCSDSAPHRMFAPKLADSLAV 672
DB 608 YNHTAA-----VELFEDLMPVYVHFH-DADGTPRTSGGGRIGTTHKMSRRLIVDSITY 660
QY 673 WTDYKIDGFREDLNGYHPRKQOILSAMERIKALNPDIYFEGSGMDSNDR-----FEIAS 728
DB 661 WVEEYKVDGFRPDMGHDIAETIOAYBRAKELNINYMIGBMTTYVGDDEIGSIMPAD 720
QY 729 QINLKGTT-GIGTFSDRLDSVYRGGPPDSGALRONOGIGSAGVLPNRLASLSDQVYVH 787
DB 721 QDMQHTESVGVFSDSEFRNELK-----SGFG-----SEGQRF 753
QY 788 LADLRLMGAGLADFVVIDKGAAGKSGE--DYNGAPGGI-AADPTEVNVYVSKHNO 844
DB 754 IT-----GAGRNIEQLFDNLTLAOPNHTATVPDVPYIEAHNL 793
QY 845 TLMDMISYKASQ---BADLATRYMNOAVSLATVMIGGIAFDQSGSEILRSKS----- 894
DB 794 TLHDVYIASIRKDPREYHQBEEIHKRIHL---GNTVVLTAQGTAFHLAAGEFGRITKFRADT 850
QY 895 -----FTRDSYDSGDMFNRVYVSLQDNVYVWGPRLSDGNSN 931
DB 851 GGAQAPYKSTYMTDENGKPFVYPIYFHSYDSSDAINKKDFMGKATNK-----EK 898
QY 932 YEV--ITRVKENVATPGAEELKQMTAFYOELTELKSSPLFTLG--DGSAMMKRYD--- 983
DB 899 YPIHHLTR-----TYTAGLIELRSTNAFTHGLVAVQAVNRLLDIP 941

```

Query Match 10.7%; Score 605.5; DB 2; Length 1072;  
 Best Local Similarity 23.38; Pred. No. 6,4e-30;

Mon Apr 23 09:45:32 2001

us-09-262-126c-4.rsp

Page 12

```
OY      984 -----FRNTGSDQAGLVMYTVDGKAKASL-DSRLDGLVY 1019  
          : | : | : | : | :  
DB     942 ISETDLVGVRATSTDNCTGYFVFVNADTKERTSLTDDLEGVVI 987
```

Search completed: April 22, 2001, 09:11:12  
Job time: 373 sec